TRANSACTIONS

OF THE

AMERICAN MATHEMATICAL SOCIETY

VOLUME 53
JANUARY TO JUNE, 1943

PUBLISHED BY THE SOCIETY
MENASHA, WIS., AND NEW YORK
1943

Chester C. Corbin Library Fund
Bound June, 1943

Composed, Printed and Bound by
The Collegiate Press
George Banta Publishing Company
Menasha, Wisconsin

TABLE OF CONTENTS
VOLUME 53, JANUARY TO JUNE, 1943

1★

TORSION OF HOLLOW CYLINDERS

BY

R. C. F. BARTELS

1. Introduction. The torsion problem for the (solid) cylinder whose cross section is a simply connected region has received considerable attention in recent literature. Outstanding among the published works which emphasize methods are the investigations by: Trefftz [9] and the generalizations of his method by Seth [7]; Muschelišvili [5] and the applications of his method by Sokolnikoff and Sokolnikoff [7]; and more recently Stevenson [8] and the extension of his method by Morris [4].

The torsion problem for the (hollow) cylinder whose cross section is a doubly connected region, on the other hand, has not enjoyed such propitious attention. The present analytical methods of treating this form of the problem have been improved very little since the close of the nineteenth century when Macdonald [3] obtained a solution for the region bounded by eccentric circles making use of curvilinear orthogonal coordinates; the solution of the torsion problem for the region bounded by confocal ellipses was published by Greenhill [1] several years earlier employing the same method. It should be remarked that the experimental methods—for example, the membrane analogy which was pointed out by Prandtl [12] and later improved by Griffith and Taylor [13], and Trayer and March [14]—are readily extended to the case of multiply connected regions.

The purpose of this paper is twofold: first, to supply the need for a general method of obtaining a computable solution of boundary value problems of Dirichlet type for doubly connected regions, and second, to apply this method to obtain the solution of the torsion problem for certain hollow cylinders.

The procedure of determining the solutions of the torsion problem for the doubly connected regions considered is in each case to map the region conformally upon an annulus, and then to solve the related Dirichlet problem for the simpler region. To this end a formula for the solution of the general Dirichlet problem for the annulus is developed which, though lacking the elegance of the well known integral formula of Villat [10], lends itself readily for purposes of computation.

2. Solution of the problem for the annular region. Let γ_1 and γ_2 denote the circles $|\zeta| = r_1$, and $|\zeta| = r_2$, $r_1 < r_2$, respectively, in the plane of the complex variable ζ. Also, let the *real* functions $u_1(\sigma_1)$ and $u_2(\sigma_2)$, where $\sigma_1 = r_1e^{i\phi}$, $\sigma_2 = r_2e^{i\phi}$ (ϕ real), be periodic and continuous for all values of ϕ with period 2π and such that

Presented to the Society, April 8, 1938, under the title *Saint-Venant's flexure problem for a regular polygon*; received by the editors July 14, 1941.

$$(1)\qquad \frac{1}{2\pi i}\int_{\gamma_j}\frac{u_j(\sigma_j)}{\sigma_j}d\sigma_j=\frac{1}{2\pi}\int_0^{2\pi}u_j(\sigma_j)d\phi=A \qquad (j=1,2),$$

A representing the common value of the integrals. Then the function(1)

$$(2)\qquad f(\zeta)=\frac{1}{\pi i}\int_{\gamma_2}\frac{U_2(\sigma_2)}{\sigma_2-\zeta}d\sigma_2-\frac{1}{\pi i}\int_{\gamma_1}\frac{U_1(\sigma_1)}{\sigma_1-\zeta}d\sigma_1+\text{const.},$$

where the functions $U_1(\sigma_1)$ and $U_2(\sigma_2)$ are defined by the integral equations(2)

$$(3)\qquad u_j(\sigma_j)=U_j(\sigma_j)+\Re\left\{\frac{1}{\pi i}\int_{\gamma_k}\frac{U_k(\sigma_k)}{\sigma_2-\sigma_1}d\sigma_k\right\} \qquad (j,k=1,2;\ j\neq k),$$

is single-valued and regular for $r_1<|\zeta|<r_2$ and, except for an additive constant, its real part takes on the values $u_1(\sigma_1)$, $u_2(\sigma_2)$ on the circles γ_1, γ_2, respectively. The existence of the functions U_1, U_2 has been established for functions u_1, u_2 satisfying much more general conditions than those considered here. It is well known that the condition (1) is both necessary and sufficient in order that the function $f(\zeta)$ determined by the values $u_1(\sigma_1)$ and $u_2(\sigma_2)$ prescribed on γ_1 and γ_2, respectively, be single-valued.

If the functions u_1 and u_2 are replaced by two new functions satisfying (1) and differing from u_1 and u_2 by constants, the function $f(\zeta)$ is altered only by the addition of a constant. It can therefore be assumed, without restricting the application of the formula (2), that the constant A in (1) is zero. In this event,

$$(4)\qquad \int_0^{2\pi}U_j(\sigma_j)d\phi=0 \qquad (j=1,2).$$

In addition to the conditions given above, let $u_1(\sigma_1)$ and $u_2(\sigma_2)$ be absolutely continuous functions of ϕ in the interval $0\leqq\phi\leqq 2\pi$. Then $U_1(\sigma_1)$, $U_2(\sigma_2)$ are also absolutely continuous functions of ϕ in the same interval. Under these conditions the infinite series

$$(5)\qquad u_j(\sigma_j)=\sum_{m=1}^{\infty}\left[\overset{(j)}{a_m}\sigma_j^{m}+\overset{(j)}{a_{-m}}\sigma_j^{-m}\right],\quad U_j(\sigma_j)=\sum_{m=1}^{\infty}\left[\overset{(j)}{A_m}\sigma_j^{m}+\overset{(j)}{A_{-m}}\sigma_j^{-m}\right] \qquad (j=1,2),$$

where

$$(6)\qquad \overset{(j)}{a_m}=\frac{1}{2\pi i}\int_{\gamma_j}\frac{u_j(\sigma_j)}{\sigma_j^{m+1}}d\sigma_j,\quad \overset{(j)}{A_m}=\frac{1}{2\pi i}\int_{\gamma_j}\frac{U_j(\sigma_j)}{\sigma_j^{m+1}}d\sigma_j \qquad (j=1,2;\ m=1,2,\cdots),$$

(1) Cf. G. C. Evans, *The logarithmic potential*, Amer. Math. Soc. Colloquium Publications, vol. 6, 1927, pp. 112–117.

(2) The symbol $\Re\{F\}$ is understood to mean the real part of F.

converge uniformly for all values of ϕ; the constant terms corresponding to $m=0$ are absent from these series as a consequence of equations (1) and (4). Therefore, by equations (2) and (5), it follows that, for $r_1<|\zeta|<r_2$,

$$(7)\qquad f(\zeta) = 2\sum_{m=1}^{\infty} A_m^{(2)}\zeta^m + 2\sum_{m=1}^{\infty} A_{-m}^{(1)}\zeta^{-m} + \text{const.}$$

Also, on substituting the series (5) in (3) and equating coefficients, it is seen that

$$(8)\qquad A_m^{(2)} = \frac{r_2^{2m} a_m^{(2)} - r_1^{2m} a_m^{(1)}}{r_2^{2m} - r_1^{2m}}, \qquad A_m^{(1)} = \frac{r_2^{2m} a_{-m}^{(1)} - r_1^{2m} a_{-m}^{(2)}}{r_2^{2m} - r_1^{2m}}.$$

Since $r_1<r_2$, the coefficients in (8) can be written in the form of absolutely convergent infinite series as follows:

$$A_m^{(2)} = a_m^{(2)} + [a_m^{(2)} - a_m^{(1)}]\sum_{n=1}^{\infty} p_n^m, \qquad A_{-m}^{(1)} = a_{-m}^{(1)} + [a_{-m}^{(1)} - a_{-m}^{(2)}]\sum_{n=1}^{\infty} p_n^m,$$

where

$$(9)\qquad p_n = (r_1/r_2)^{2n} < 1.$$

Let these be substituted in the infinite series in (7), which also converges absolutely when $r_1<|\zeta|<r_2$. Then, with the aid of the first of equations (6) and equations (1) in which $A=0$, and by interchanging the order of summation with respect to m and n, it follows, after rearranging terms, that

$$(10)\qquad f(\zeta) = \frac{1}{\pi i}\sum_{n=-\infty}^{\infty}\left[\int_{\gamma_2}\frac{u_2(\sigma_2)}{\sigma_2 - p_n\zeta}\,d\sigma_2 - \int_{\gamma_1}\frac{u_1(\sigma_1)}{\sigma_1 - p_n\zeta}\,d\sigma_1\right] + \text{const.}$$

whenever $r_1<|\zeta|<r_2$.

The regularity of the function $f(\zeta)$ defined by the infinite series in (10) is easily established under less restrictive conditions on the functions $u_1(\sigma_1)$ and $u_2(\sigma_2)$ with the aid of the following interesting lemma:

LEMMA. *Let $F(z)$ be regular and $|F(z)|<M$ for $|z|<r$, and let $F(0)=0$. Then the infinite series $\sum_{n=1}^{\infty}F(q^nz)$, where $0<q<1$, converges uniformly and absolutely for $|z|\leqq r$ and, consequently, defines an analytic function which is regular for $|z|\leqq r$. If $F(z)$ is regular and $|F(z)|<M$ for $|z|>r$, and if $F(\infty)=0$, then the same conclusions hold for the infinite series $\sum_{n=1}^{\infty}F(q^nz)$, where $q>1$, when $|z|\geqq r$.*

The lemma follows at once from the lemma of Schwarz([3]). For, applying the latter,

([3]) Cf. E. F. Titchmarsh, *The theory of functions*, 2d edition, London, 1939, p. 168.

$$\sum_{n=1}^{\infty}|F(q^n z)| \leqq \sum_{n=1}^{\infty} \frac{M|z|}{r} q^n \leqq M\sum_{n=1}^{\infty} q^n = \frac{Mq}{1-q}$$

whenever $|z| \leqq r$. Hence the series converges absolutely and uniformly if $|z| \leqq r$. The second part of the lemma is proved in like manner.

Equation (10) can be written in the form

$$(11)\qquad f(\zeta) = \frac{1}{\pi i}\int_{\gamma_2} \frac{u_2(\sigma_2)}{\sigma_2 - \zeta} d\sigma_2 - \frac{1}{\pi i}\int_{\gamma_1} \frac{u_1(\sigma_1)}{\sigma_1 - \zeta} d\sigma_1 + \Sigma(\zeta) + \text{const.},$$

where $\Sigma(\zeta)$ represents the infinite series obtained from that in (10) by omitting the term corresponding to $n=0$. The application of the lemma to the infinite series $\Sigma(\zeta)$ is at once evident with the aid of the following inequalities which, in view of the inequality (9), are seen to hold for $n \geqq 1$ and $r_1 \leqq |\zeta| \leqq r_2$:

$$(12)\qquad |p_{-n}| > 1, \qquad |p_n\zeta| < r_1, \qquad |p_{-n}\zeta| > r_2.$$

If $u_1(\sigma_1)$ and $u_2(\sigma_2)$ are merely bounded and integrable functions of ϕ in the interval $(0, 2\pi)$, the terms of the infinite series $\Sigma(\zeta)$ for $n \geqq 1$ are bounded and regular when $|\zeta| \leqq r_2$ and, as a consequence of equations (1) with $A=0$, vanish for $\zeta=0$. The terms corresponding to $n \leqq -1$ are bounded and regular for $|\zeta| \geqq r_1$ and are seen to vanish for $\zeta=\infty$. Therefore by the lemma the series $\Sigma(\zeta)$ converges absolutely and uniformly for $r_1 \leqq |\zeta| \leqq r_2$. Hence the function $\Sigma(\zeta)$ is regular for $r_1 < |\zeta| < r_2$ and continuous on the circles γ_1 and γ_2.

It will now be shown that the real part of the function $f(\zeta)$ defined in the equation (10) takes on appropriate values on the circles γ_1 and γ_2. For, if $\zeta_2 = r_2 e^{i\theta}$ represents an arbitrary value of ζ on γ_2, it is readily verified using equation (1) with $A=0$ and elementary manipulations that

$$\begin{aligned}\overline{\Sigma(\zeta_2)} &= -\Sigma(\zeta_2) + \frac{1}{\pi i}\int_{\gamma_1} \frac{u_1(\sigma_1)}{\sigma_1 - \zeta_2} d\sigma_1 - \frac{1}{\pi i}\int_{\gamma_1} \frac{u_1(\sigma_1)}{\sigma_1 - p_1\zeta_2} d\sigma_1,\\ &= -\Sigma(\zeta_2) + 2\Re\left\{\frac{1}{\pi i}\int_{\gamma_1} \frac{u_1(\sigma_1)}{\sigma_1 - \zeta_2} d\sigma_1\right\},\end{aligned}$$

where the bar denotes the conjugate values. Therefore

$$\Re\{\Sigma(\zeta_2)\} = \Re\left\{\frac{1}{\pi i}\int_{\gamma_1} \frac{u_1(\sigma_1)}{\sigma_1 - \zeta_2} d\sigma_1\right\}.$$

As a consequence of this, it follows from (11) that

$$\lim_{\zeta\to\zeta_2} \Re\{f(\zeta)\} = \lim_{\zeta\to\zeta_2} \Re\left\{\frac{1}{\pi i}\int_{\gamma_2} \frac{u_2(\sigma_2)}{\sigma_2 - \zeta} d\sigma_2\right\} + \text{real const.}$$

Under the assumption that $u_2(\sigma_2)$ be bounded and integrable, the limit on the

right-hand side of the last equation exists and is equal to $u_2(\sigma_2)$ if ζ_2 is a point of continuity of $u_2(\sigma_2)$[4].

The behavior of the function $f(\zeta)$ on the circle γ_1 can be studied in like manner. The results of the foregoing investigation can be stated as follows[5]: *Let the functions* $u_1(\sigma_1)$ *and* $u_2(\sigma_2)$, *satisfying* (1) *with* $A=0$, *be bounded and integrable with respect to* ϕ *in the interval* $(0, 2\pi)$. *Then the analytic function* $f(\zeta)$ *which is single-valued and regular in the annular region between the circles* γ_1 *and* γ_2 *and whose real part takes on the values* $u_1(\sigma_1)$ *and* $u_2(\sigma_2)$ *at all points of* γ_1 *and* γ_2, *respectively, at which these functions are continuous is given by the formula in equation* (10).

Formula (10) will be seen to form the basis for a very general method of treating a class of boundary value problems related to doubly connected regions. In the following sections this formula is applied to the problem of the torsion of hollow cylinders.

3. **Statement of the torsion problem.** In the theory of elasticity the St. Venant's torsion problem for a region D of the xy-plane may be formulated as that of determining an analytic function $F(z)$ of the complex variable $z=x+iy$ which is single-valued and regular within D and such that at points of its boundary C[6]

$$2\Re\{F(z)\} = x^2 + y^2 + \text{const.} \tag{13}$$

Given the function $F(z)$, the important physical quantities for a twisted, homogeneous, cylindrical beam whose cross section has the shape of the region D can be determined. For example, the well known formulas for the shearing stresses X_z and Y_z at a point in any cross section can be written in the form

$$Y_z + iX_z = \tau G[z - F'(z)],$$

where τ is the "twist," and G is the slide modulus of the material constituting the beam. Its torsional rigidity J is readily written in the form

$$J = GI_0 - \frac{iG}{4}\int_C [F(z) - \overline{F(z)}]d(x^2 + y^2),$$

where I_0 is the moment of inertia of the region D with respect to the origin, and the integral is taken over the complete boundary C. Making use of (13) and the fact that $F(z)$ is single-valued, the latter can be written

$$J = GI_0 + \frac{G}{2}\Re\left\{ i\int_C \overline{F(z)}dF(z)\right\}.$$

---

(4) Evans, loc. cit., pp. 39, 65.
(5) Compare with Villat [10].
(6) Cf. Love [2, p. 314].

In the present paper D is taken as a doubly connected region bounded internally and externally by the closed Jordan curves C_1 and C_2, respectively. In view of the significance of equation (1) of the preceding section, it is evident that the values of the constants in (13) on the curves C_1 and C_2 are not independent; otherwise the function $F(z)$ determined by the boundary condition is not necessarily single-valued in D. However, apart from this restriction, the constants are arbitrary.

Let the function $z=\omega(\zeta)$ map D one-one and conformally on the interior of an annular region bounded by two circles $|\zeta|=r_1$ and $|\zeta|=r_2$, $r_1<r_2$, in the plane of the complex variable ζ. As before, these circles are denoted by γ_1 and γ_2 and the values of ζ on them by $\sigma_1=r_1e^{i\phi}$ and $\sigma_2=r_2e^{i\phi}$, respectively. It is understood, of course, that the radii of the circles γ_1 and γ_2 are not independent, the ratio r_1/r_2 being determined uniquely by the region D(7). Moreover the mapping is known to be topological on the boundary, so that the function $\omega(\zeta)$ is continuous on γ_1 and γ_2.

Given the mapping function $z=\omega(\zeta)$, the torsion problem for the doubly connected region D can be transformed into a corresponding boundary value problem for an annulus whose solution is given in the foregoing section. For, if $f(\zeta)$ represents the values of $F(z)$ in the ζ-plane, that is, $f(\zeta)=F[\omega(\zeta)]$, then $f(\zeta)$ is single-valued and regular for $r_1<|\zeta|<r_2$ and, according to (13), its real part satisfies the conditions

$$\Re\{f(\sigma_j)\} = u_j(\sigma_j) = \frac{\omega(\sigma_j)\overline{\omega(\sigma_j)}}{2} + c_j \qquad (j = 1, 2), \tag{14}$$

where the constants c_1, c_2 are taken so that equations (1) with $A=0$ are satisfied. Since $\omega(\zeta)$ is continuous on γ_1 and γ_2, the functions $u_1(\sigma_1)$, $u_2(\sigma_2)$ are continuous on γ_1, γ_2, respectively. Thus the function $f(z)$ satisfies the conditions of the theorem stated at the close of the preceding section and, therefore, is given by the formula (10).

The function $F(z)$ representing the solution of the torsion problem for the region D is obtained from $f(\zeta)$ by making the inverse of the transformation $z=\omega(\zeta)$. On the other hand, this process of inversion is more or less superfluous since the important physical quantities for the cylindrical beam with cross section D can easily be expressed in terms of the function $f(\zeta)$. Thus the shearing stress at a point z of the cross section is given by

$$Y_z + iX_z = \tau G\left[\overline{\omega(\zeta)} - \frac{f'(\zeta)}{\omega'(\zeta)}\right],$$

where ζ is the point of the annular region in the ζ-plane corresponding to z

(7) For particulars on the mapping of multiply connected regions, see C. Carathéodory, *Conformal representations*, Cambridge Tracts in Mathematics and Mathematical Physics, no. 28, London, 1932, pp. 70–73.

under the transformation $z=\omega(\zeta)$. Moreover the torsional rigidity J is given by

$$J = GI_0 + \frac{G}{2}\Re\left\{ i \int \overline{f(\zeta)}df(\zeta) \right\},$$

where the integral is taken over the complete boundary of the annulus in the positive sense.

4. Torsion of the eccentric ring. Let

$$z = \omega(\zeta) \equiv \frac{\zeta + a}{1 + a\zeta}, \tag{15}$$

where $a<1$. Then the region D of the z-plane corresponding to the annular region $r<|\zeta|<1$ of the ζ-plane is that which is bounded internally and externally by the circles $|z-c|=R$ and $|z|=1$, respectively, where

$$c = \frac{a(1 - r^2)}{1 - a^2r^2}, \qquad R = \frac{r(1 - a^2)}{1 - a^2r^2}.$$

This mapping of the region D upon the annulus is such that the points of the circles $\gamma_1(|\zeta|=r)$ and $\gamma_2(|\zeta|=1)$ correspond, respectively, to points of the circles $|z|=1$ and $|z-c|=R$.

Since in this case $\sigma_2\bar{\sigma}_2=1$, it follows that $\omega(\sigma_2)\overline{\omega(\sigma_2)}=1$. Also, by the theory of residues, it is seen that

$$\int_{\gamma_1} \frac{\omega(\sigma_1)\overline{\omega(\sigma_1)}}{\sigma_1} d\sigma_1 = 2\pi i(1 - h),$$

where

$$h = \frac{(1 - a^2)(1 - r^2)}{(1 - a^2r^2)}. \tag{16}$$

Accordingly, the functions $u_1(\sigma_1)$ and $u_2(\sigma_2)$ in the boundary condition (14) are given by

$$u_1(\sigma_1) \equiv [\omega(\sigma_1)\overline{\omega(\sigma_1)} + h - 1]/2, \qquad u_2(\sigma_2) \equiv 0. \tag{17}$$

It follows from the second of equations (17) that the first integral of each term of the infinite series in (10) vanishes. Further, with due regard for the inequalities in (12) in which $p_n=r^{2n}$, a simple evaluation of residues gives, for $r<|\zeta|<1$,

$$\int_{\gamma_1} \frac{u_1(\sigma_1)}{\sigma_1 - p_n\zeta} d\sigma_1 = \begin{cases} \dfrac{ahp_n\zeta}{1 + ap_n\zeta} & \text{when } n > 0, \\[2ex] -\dfrac{ah}{a + p_{n-1}\zeta} & \text{when } n \leqq 0. \end{cases}$$

Therefore, in accordance with (10), the function $f(\zeta)$ which is regular for $r<|\zeta|<1$ and satisfies the condition (14) on the circles γ_1 and γ_2 can be written in the form

$$f(\zeta) = h[L(\zeta) - L(1/\zeta)] + \text{const.}, \tag{18}$$

where,

$$L(\zeta) = \sum_{n=1}^{\infty} \frac{ar^{2n}}{\zeta + ar^{2n}}.$$

By a transformation(8) of the Lambert series $L(\zeta)$ in (18), the function $f(\zeta)$ can also be written in the form(9)

$$f(\zeta) = \frac{ahr^2}{ar^2 + \zeta} + h[P(\zeta/r) - P(r/\zeta)] + \text{const.} \quad (r < |\zeta| < 1), \tag{19}$$

where

$$P(t) = \sum_{n=1}^{\infty} (-1)^n \frac{a^n r^{3n}}{1 - r^{2n}} t^n.$$

The latter form has the advantage over the former from the point of view of computation.

5. **Torsion of a hollow lune.** Let

$$z = \omega(\zeta) \equiv \frac{1 - (1 - \zeta^2)^{1/2}}{\zeta} \qquad (|\zeta| < 1), \tag{20}$$

where that branch of the square root is chosen which has the value $+1$ when $\zeta = 0$. Then the annulus $r<|\zeta|<1$ of the ζ-plane corresponds to the region D of the z-plane which is bounded externally by arcs of equal radii intersecting at right angles in the points $z = \pm 1$, and internally by the oval

$$r^2[(x^2 + y^2 - 1)^2 - 4y^2] = 4(1 - r^2)(x^2 + y^2).$$

The region D is an approximation of the cross section of a common type of hollow strut used in aircraft construction. By choosing values of r sufficiently near unity, the comparison can be extended to thin cylindrical shells whose sections are in the shape of the two intersecting circular arcs determining the external boundary of D.

If as in the preceding section $\sigma_1 = re^{i\phi}$, it follows from (20) and the theory of residues that, for $r>0$,

(8) Cf. K. Knopp, *Theory and application of infinite series*, English translation by R. C. Young, London, 1928, p. 452.

(9) The form of the solution given by (19) corresponds to that which was obtained by Weinel with the aid of dipolar coordinates; see [11, p. 70]. The method employed in the present paper is certainly more direct than that used by Weinel. Macdonald's [3] form of the solution can be found in Love [2, p. 320].

$$(21)\qquad \frac{r^2}{2\pi i}\int_{\gamma_1}\frac{\omega(\sigma_1)\overline{\omega(\sigma_1)}}{\sigma_1-\mu}\,d\sigma_1=\frac{1}{\pi}J(\mu)+\begin{cases}-(1-\mu^2)^{1/2} & \text{when } |\mu|<r,\\ (1-r^4/\mu^2)^{1/2}-1 & \text{when } |\mu|>r,\end{cases}$$

where

$$J(\mu)=\frac{1}{2i}\int_{\gamma_1}\frac{((1-\sigma_1^2)(1-\bar{\sigma}_1^2))^{1/2}}{\sigma_1-\mu}\,d\sigma_1;$$

these integrals are taken around the circle γ_1 ($|\zeta|=r$). The integral $J(\mu)$ can evidently be written in the form

$$(22)\qquad J(\mu)=r(1+r^2)\int_0^{\pi}\frac{r-\mu\cos\phi}{(r^2+\mu^2)-2r\mu\cos\phi}(1-k_1^2\cos^2\phi)^{1/2}\,d\phi,$$

where

$$k_1=\frac{2r}{1+r^2}.$$

Therefore[10], if $0<r<1$,

$$(23)\qquad \begin{aligned} J(\mu)&=(1+r^2)E_1+(1+r^2)\frac{r^2-\mu^2}{r^2+\mu^2}K_1\\ &+\begin{cases}-K_1\dfrac{((1-\mu^2)(r^4-\mu^2))^{1/2}}{\mu}Z(\alpha,k_1) & \text{for } |\mu|<r,\\ K_1\dfrac{((\mu^2-1)(\mu^2-r^4))^{1/2}}{\mu}Z(\alpha,k_1) & \text{for } |\mu|>r,\end{cases}\end{aligned}$$

where K_1 and E_1 are the complete elliptic integrals of the first and second kinds, respectively, α is the elliptic integral of the first kind defined as follows

$$(24)\qquad \operatorname{sn}(\alpha,k_1)=\frac{\mu(1+r^2)}{r^2+\mu^2},$$

and $Z(\alpha, k_1)$ is the Jacobi zeta-function; these elliptic integrals all have the modulus k_1.

In particular, if $\mu=0$, equations (21) and (22) give

$$\frac{1}{2\pi i}\int_{\gamma_1}\frac{\omega(\sigma_1)\overline{\omega(\sigma_1)}}{\sigma_1}\,d\sigma_1=\frac{1}{\pi r^2}\left[2(1+r^2)E_1-\pi\right],$$

and consequently the constant c_1 in the definition of the function $u_1(\sigma_1)$ in (14) becomes

---

[10] For particulars on the evaluation of the integral $J(\mu)$, see E. T. Whittaker and G. N. Watson, *Modern analysis*, 4th edition, London, 1935, pp. 499, 518, 522.

$$c_1 = \frac{1}{2\pi r^2}[\pi - 2(1+r^2)E_1]. \tag{25}$$

Making use of an elementary quadratic transformation[11], the complete elliptic integrals K_1 and E_1 with modulus $k_1 = 2r/1+r^2$ can be expressed in terms of the complete elliptic integrals K and E with modulus r^2 as follows

$$K_1 = (1+r^2)K \quad \text{and} \quad E_1 = \frac{2E}{1+r^2} - (1-r^2)K.$$

Also by the same transformation, for $|\mu| < r$,

$$Z(\alpha, k_1) = \frac{2}{1+r^2}\left[Z(\beta, r^2) + \frac{\mu((1-\mu^2)(r^4-\mu^2))^{1/2}}{r^2+\mu^2}\right],$$

where

$$\operatorname{sn}(\beta, r^2) = \mu/r^2;$$

whereas for $|\mu| > r$

$$Z(\alpha, k_1) = \frac{2}{1+r^2}\left[Z(\beta', r^2) + \frac{r^2((\mu^2-1)(\mu^2-r^4))^{1/2}}{\mu(\mu^2+r^2)}\right],$$

where

$$\operatorname{sn}(\beta', r^2) = 1/\mu.$$

Consequently these equations, together with (14), (21), (23), and (25), give

$$\frac{r^2}{\pi i}\int_{\gamma_1}\frac{u_1(\sigma_1)}{\sigma_1-\mu}\,d\sigma_1 = \begin{cases} M-(1-\mu^2)^{1/2}-\dfrac{2K}{\pi}\,\dfrac{((1-\mu^2)(r^4-\mu^2))^{1/2}}{\mu}Z(\beta+K, r^2) & \text{for } |\mu| < r, \\[2ex] -M+(1-r^4/\mu^2)^{1/2}+\dfrac{2K}{\pi}\,\dfrac{((\mu^2-1)(\mu^2-r^4))^{1/2}}{\mu}Z(\beta'+K, r^2) & \text{for } |\mu| > r, \end{cases} \tag{26}$$

where

$$M = \frac{1}{\pi}[\pi + 2(1-r^4)K - 2E].$$

In view of the inequalities in (12) in which $p_n = r^{2n}$, equations (26) give, for $r < |\zeta| < 1$,

---

([11]) Cf. H. Hancock, *Theory of elliptic functions*, vol. 1, New York, 1910, p. 250.

$$\frac{r^2}{\pi i}\int_{\gamma_1}\frac{u_1(\sigma_1)}{\sigma_1-p_n\zeta}\,d\sigma_1$$

$$(27)\qquad =\begin{cases} M-(1-p_n^2\zeta^2)^{1/2}-\dfrac{2K}{\pi}\dfrac{((1-p_{n-1}^2\zeta^2)(1-p_n^2\zeta^2))^{1/2}}{p_{n-1}\zeta}Z(\beta_n+K,r^2) & \text{when } n\geqq 1,\\ -M+(1-p_{1-n}^2/\zeta^2)^{1/2}+\dfrac{2K}{\pi}\dfrac{((1-p_{-n}^2/\zeta^2)(1-p_{1-n}^2/\zeta^2))^{1/2}}{p_{-n}/\zeta}Z(\beta'_{-n}+K,r^2) & \text{when } n<1,\end{cases}$$

where

$$\mathrm{sn}\,(\beta_n,r^2)=p_{n-1}\zeta \quad\text{and}\quad \mathrm{sn}\,(\beta'_n,r^2)=p_n/\zeta.$$

If $\sigma_2=e^{i\phi}$ and γ_2 represents the circle $|\zeta|=1$, it follows easily from (21) and (22), on setting $r=1$, that

$$\frac{1}{2\pi i}\int_{\gamma_2}\frac{\omega(\sigma_2)\overline{\omega(\sigma_2)}}{\sigma_2-\mu}\,d\sigma_2=\frac{2}{\pi}+\frac{1-\mu^2}{2\pi\mu}\log\left(\frac{1+\mu}{1-\mu}\right)^2+\begin{cases}-(1-\mu^2)^{1/2} & \text{for } |\mu|<1,\\ (1-1/\mu^2)^{1/2}-1 & \text{for } |\mu|>1.\end{cases}$$

The constant c_2 in the definition of the function $u_2(\sigma_2)$ in (14) is therefore

$$c_2=\frac{1}{2}-\frac{2}{\pi}\cdot$$

Consequently these equations, together with the inequalities in (12), give, for $r<|\zeta|<1$,

$$\frac{1}{\pi i}\int_{\gamma_2}\frac{u_2(\sigma_2)}{\sigma_2-p_n\zeta}\,d\sigma_2$$

$$(28)\qquad =\begin{cases} 1-\dfrac{2}{\pi}-(1-p_n^2\zeta^2)^{1/2}+\dfrac{1-p_n^2\zeta^2}{\pi p_n\zeta}\log\dfrac{1+p_n\zeta}{1-p_n\zeta} & \text{for } n\geqq 0,\\ -1+\dfrac{2}{\pi}+(1-p_{-n}^2/\zeta^2)^{1/2}-\dfrac{1-p_{-n}^2/\zeta^2}{\pi p_{-n}/\zeta}\log\dfrac{1+p_{-n}/\zeta}{1-p_{-n}/\zeta} & \text{for } n>0.\end{cases}$$

Thus, by equations (10), (27), and (28), the function $f(\zeta)$ which is regular for $r<|\zeta|<1$ and satisfies the condition (14) on the circles γ_1 and γ_2 can, after proper rearrangement of terms, be written in the form

$$(29)\qquad \begin{aligned} f(\zeta)=&-(1-\zeta^2)^{1/2}+\frac{1-\zeta^2}{\pi\zeta}\log\frac{1+\zeta}{1-\zeta}\\ &+\sum_{n=1}^{\infty}[T_n(\zeta)-T_n(1/\zeta)]+\text{const.},\end{aligned}$$

where

$$T_n(\zeta) = \frac{1 - r^2}{r}(1 - r^{4n}\zeta^2)^{1/2} + \frac{1 - r^{4n}\zeta^2}{\pi r^{2n}\zeta}\left[\log\frac{1 + r^{2n}\zeta}{1 - r^{2n}\zeta} + 2K\left(\frac{1 - r^{4n-4}\zeta^2}{1 - r^{4n}\zeta^2}\right)^{1/2} Z(\beta_n + K, r^2)\right]$$

and

$$\operatorname{sn}(\beta_n, r^2) = r^{2n-2}\zeta.$$

Since $\operatorname{sn}(0, r^2) = 0$ and $Z(K, r^2) = 0$, it is at once evident that

$$\lim_{r \to 0}\left[T_n(\zeta) - T_n(1/\zeta)\right] = 0$$

for $n \geqq 1$. Consequently as r tends to zero the function $f(\zeta)$ defined in (29) reduces to the solution of the torsion problem for the simply connected region bounded by two circular arcs of equal radii and intersecting at right angles at the points $z = \pm 1$(12).

Bibliography

1. A. G. Greenhill, *Fluid motion between confocal elliptic cylinders and confocal ellipsoids*, Quart. J. Math. Oxford Ser. vol. 16 (1879) pp. 227–256.

2. A. E. H. Love, *A treatise on the mathematical theory of elasticity*, 4th edition, London, 1927, pp. 310–328.

3. H. M. Macdonald, *On the torsional strength of a hollow shaft*, Proc. Cambridge Philos. Soc. vol. 8 (1893) pp. 62–68.

4. R. M. Mooris, *The internal problems of two dimensional potential theory*, Math. Ann. vol. 117 (1939) pp. 31–38; and *Some general solutions of St. Venant's flexure and torsion problem*. I, Proc. London Math. Soc. (2) vol. 46 (1940) pp. 81–98.

5. N. Muschelisvilli, *Sur le problème de torsion des cylindres elastique isotropes*, Rendiconti della Reale Accademia die Lincei (6) vol. 9 (1929) pp. 295–300.

6 B. R. Seth, *On the general solution of a class of physical problems*, Philosophical Magazine (7) vol. 20 (1935) pp. 632–640.

7. E. S. Sokolnikoff and I. S. Sokolnikoff, *Torsion of regions bounded by circular arcs*, Bull. Amer. Math. Soc. vol. 44 (1938) pp. 384–387.

8. A. C. Stevenson, *Flexure with shear and associated torsion in prisms of uni-axial and asymmetric cross-sections*, Transactions of the Royal Society of London (A) vol. 237 (1938) pp. 161–229.

9. E. Trefftz, *Über die Torsion prismatische Stäbe von polygonalem Querschnitt*, Math. Ann. vol. 82 (1921) pp. 97–112.

10. H. Villat, *Le problème de Dirichlet dans une aire annulaire*, Rend. Circ. Mat. Palermo vol. 33 (1912) pp. 134–173; p. 147.

11. E. Weinel, *Das Torsionsproblem für den exzentrischen Kriesring*, Ingenieur Archiv vol. 3 (1932) pp. 67–75.

12. L. Prandtl, *Zur Torsion von prismatischen Stäben*, Physikalische Zeitschrift, vol. 4 (1903) pp. 758–770.

---

(12) Cf. Sokolnikoff and Soknolnikoff [7, p. 386].

13. A. A. Griffith and G. I. Taylor, *The use of soap films in solving torsion problems*, Reports and Memoranda, (n.s.), no. 333, June 1917. Technical Report of the Advisory Committee for Aeronautics (British) for 1917–1918, vol. 3.

14. G. W. Trayer and H. W. March, *The torsion of members having sections common in aircraft construction*, National Advisory Committee for Aeronautics, report no. 334, 1930.

The University of Michigan,
Ann Arbor, Mich.

HEAT CONDUCTION IN AN INFINITE COMPOSITE SOLID WITH AN INTERFACE RESISTANCE

BY

W. A. MERSMAN

1. **Introduction.** The purpose of this paper is to solve the problem of one-dimensional heat conduction in a body composed of two plane-boundary semi-infinite homogeneous solids of different materials in "imperfect contact" along their interface. ("Imperfect contact" is defined by equation (4) below.) The Laplace transformation method is used both in discovering and in rigorously establishing the solution, by means of the inversion theorems of Churchill and Doetsch(1). The question of uniqueness is not considered here.

2. **The boundary value problem.** Let t denote time, x the perpendicular distance from the interface, a_ν and k_ν the thermal diffusivities and conductivities, respectively, of the two materials, and λ the interface resistance(2). Throughout, $f(x)$ is a known function, integrable over any finite interval, such that

$$|f(x)| \leqq \alpha \exp(\beta |x|),$$

where α and β are non-negative constants. If $U(x, t)$ is the temperature, we have the following boundary value problem(3):

$$(1)\qquad \frac{\partial U}{\partial t} = a_\nu \frac{\partial^2 U}{\partial x^2}, \qquad t > 0,\ x \neq 0,$$

$$(2)\qquad \lim_{t\to 0} U(x, t) = f(x), \qquad x \neq 0,$$

$$(3)\qquad \lim_{x\to -0} k_1 \frac{\partial U}{\partial x} = \lim_{x\to +0} k_2 \frac{\partial U}{\partial x}, \qquad t > 0,$$

$$(4)\qquad U(+0, t) - U(-0, t) = \lim_{x\to 0} \lambda k_\nu \frac{\partial U}{\partial x}, \qquad t > 0.$$

The constants, a_ν, k_ν, and λ are all assumed to be greater than zero.

Presented to the Society, April 11, 1942; received by the editors November 22, 1941.

(1) Cf. Churchill, *The solution of linear boundary-value problems in physics by means of the Laplace transformation*. Part I. *A theory for establishing a solution in the form of an integral, for problems with vanishing initial conditions*, Math. Ann. vol. 114 (1937) pp. 591–613. This paper will be designated by [C].

Cf. Doetsch, *Theorie und Anwendung der Laplace Transformation*, Berlin, 1937. This book will be designated by [D].

(2) Cf. Riemann-Weber, *Die Partiellen Differential-Gleichungen der Mathematischen Physik*, 5th edition, 1912, vol. 2 pp. 85, 100.

(3) Throughout, $\nu=1$ if $x<0$, $\nu=2$ if $x>0$.

Define $\Phi_\nu(t)$ as follows:

(5) $$\Phi_1(t) = U(-0, t), \quad \Phi_2(t) = U(+0, t), \qquad t > 0.$$

Then the solution can be written by means of Φ in the well known form([4])

(6) $$U(x, t) = V_\nu(x, t) + \int_0^t \Phi_\nu(t - \xi)\Psi(xa_\nu^{-1/2}, \xi)d\xi, \qquad t > 0,$$

where

(7) $$\Psi(x, t) = x \exp(-x^2/4t)/2(\pi t^3)^{1/2},$$

(8) $$V_1(x, t) = \frac{1}{2}\int_{-\infty}^0 f(\xi)\{\Theta(x - \xi a_1 t) - \Theta(x + \xi a_1 t)\}d\xi, \qquad t > 0,$$

(9) $$\Theta(x, t) = (\pi t)^{-1/2} \exp(-x^2/4t),$$

and V_2 is obtained from V_1 upon replacing a_1 by a_2 and integrating from 0 to $+\infty$.

Equations (3)–(6) now give two simultaneous integral equations for the unknown functions $\Phi_\nu(t)$. These will be solved by the method of the Laplace transformation.

3. The transformed problem and its solution. Throughout we use the singly-infinite Laplace transformation:

$$L\{U(x, t)\} = \int_0^\infty e^{-st}U(x, t)dt = u(x, s).$$

For the sake of definiteness let s be any complex number whose real part is greater than $32\beta^2$. Denoting the transforms of U, V, and so on, by u, v, and so on, respectively, the transforms of equations (3), (4), (6), are:

(3′) $$\lim_{x\to -0} k_1 \frac{\partial u(x, s)}{\partial x} = \lim_{x\to +0} k_2 \frac{\partial u(x, s)}{\partial x},$$

(4′) $$\phi_2(s) - \phi_1(s) = \lim_{x\to 0} \lambda k_\nu \frac{\partial u(x, s)}{\partial x},$$

(6′)([5]) $$u(x, s) = \phi_\nu(s) \exp[-|x|(s/a_\nu)^{1/2}] + v_\nu(x, s).$$

Eliminating ϕ_ν from these equations we obtain the transformed solution:

(10) $$u(x, s) = y_\nu(x, s) + w_\nu(x, s)$$

([4]) Cf. H. S. Carslaw, *The mathematical theory of the conduction of heat in solids*, 2d edition, London, 1921, §§18, 23.

([5]) The first term of the right member is obtained by the "Faltung" rule. Cf. [D, chap. 8]. For the specific transformations used, see the table in [D, Appendix 2].

where

$$
\begin{aligned}
y_1(x, s) = {} & [k_2/a_2^{1/2}(A + Bs^{1/2})]\Big\{(a_2 s)^{-1/2}\int_0^{\infty} f(\xi) \\
& \cdot \exp\left[- s^{1/2}\left| \xi a_2^{-1/2} - x a_1^{-1/2}\right| \right]d\xi \\
& - (a_1 s)^{-1/2}\int_{-\infty}^{0} f(\xi) \exp\left[-\left| x + \xi\right| (s/a_1)^{1/2}\right]d\xi\Big\}, \qquad x < 0,
\end{aligned}
\tag{11}
$$

$$
A = k_1 a_1^{-1/2} + k_2 a_2^{-1/2}, \qquad B = \lambda k_1 k_2 (a_1 a_2)^{-1/2},
\tag{12}
$$

$$
\begin{aligned}
w_1(x, s) = {} & [1/2(a_1 s)^{1/2}]\int_{-\infty}^{0} f(\xi)\big\{\exp\left[-\left| x - \xi\right| (s/a_1)^{1/2}\right] \\
& + \exp\left[-\left| x + \xi\right| (s/a_1)^{1/2}\right]\big\}d\xi, \qquad x < 0,
\end{aligned}
\tag{13}
$$

and y_2, w_2 are obtained from y_1, w_1, respectively, upon interchanging the subscripts 1 and 2 and the integrals $\int_0^{\infty}$ and $\int_{-\infty}^{0}$.

4. **The inversion problem.** It is easily seen that the operations $\partial/\partial x$ and $\lim_{s\to 0}$ are interchangeable with the integrations in (11) and (13), and hence that (10) is an actual solution of the boundary value problem (3′) and (4′). To show that the inverse Laplace transform of $u(x, s)$ exists and is a solution of the original boundary value problem, we use the inversion theorems of Doetsch and Churchill. For the inverse we use the notation:

$$
L^{-1}\{u(x, s)\} = \lim_{\omega\to\infty} \frac{1}{2\pi i}\int_{\gamma-\omega i}^{\gamma+\omega i} e^{ts} u(x, s)ds;
$$

that is, L^{-1} always means this particular form of the inverse, whether or not it is true that $u(x, s) = L\{L^{-1}[u(x, s)]\}$. Throughout, $\gamma > 32\beta^2$, $R(s^{1/2}) \geqq 0$.

Note that y and w are composed of two types of integrals, according as x and ξ have the same or opposite signs. It is sufficiently general to consider the following:

$$
z(x, s) = s^{-1}\int_0^{\infty} f(\xi) \exp\left[- (x + \xi)s^{1/2}\right]d\xi, \qquad x \geqq 0,
\tag{14}
$$

$$
g(x, s) = s^{-1/2}\int_0^{\infty} f(\xi) \exp\left[-\left| x - \xi\right| s^{1/2}\right]d\xi, \qquad x \geqq 0.
\tag{15}
$$

We begin with some inequalities. Noting that $|e^{\zeta}| = \exp[R(\zeta)]$ and that $R(s^{1/2}) = [(R(s) + |s|)/2]^{1/2}$, the following are easily obtained:

$$
\left| s^n z(x, s)\right| \leqq 2^{3/2}\alpha \left| s\right|^{n-3/2} \exp\left[- x(\left| s\right| /2)^{1/2}\right], \qquad n \geqq 0.
\tag{16}
$$

$$
\left|\frac{\partial^n g(x, s)}{\partial x^n}\right| \leqq 3\alpha(2)^{1/2} e^{\beta x}\left| s\right|^{-1+n/2}, \qquad n = 0, 1.
\tag{17}
$$

$$
\begin{aligned}
&\int_{\xi_1}^{\infty} \left| s^{(n-1)/2} \exp \left[ts - |\xi - x| s^{1/2}\right] f(\xi) \right| d\xi \\
(18) \qquad &\qquad \leqq 2^{3/2}\alpha |s|^{-1+n/2} \exp \left[t\gamma - \xi_1(|s|/2)^{1/2}/4\right], \\
&\qquad \gamma = R(s), \quad \xi_1 > 4x \geqq 0, \quad n = 0, 1.
\end{aligned}
$$

By means of these inequalities the following lemmas can be proved:

LEMMA 1. *If* $0<x_1\leqq x\leqq x_2$, $0\leqq t\leqq t_1$, $n=0, 1$, *then* $L^{-1}\{s^n z(x, s)\}$ *and* $L^{-1}\{\partial^2 z(x, s)/\partial x^2\}$ *exist and converge uniformly in* x *and* t.

This follows immediately from inequality (16).

LEMMA 2. *If* $0\leqq x\leqq x_1$, $0<t_1\leqq t\leqq t_2$, *then*

$$L^{-1}\{g(x, s)\}$$

exists and converges uniformly in x *and* t.

Proof. We wish to show that, given any $\epsilon>0$, there exists an $\Omega(\epsilon)$, independent of x and t such that, if $\omega_2>\omega_1>\Omega$, then

$$\left| \int_{\gamma-\omega_2 i}^{\gamma-\omega_1 i} e^{ts} g(x, s)ds + \int_{\gamma+\omega_1 i}^{\gamma+\omega_2 i} e^{ts} g(x, s)ds \right| < \epsilon.$$

By equation (15) and inequality (18), the order of integration can be inverted. Hence, consider

$$
\begin{aligned}
\Bigg| &\int_0^{\infty} f(\xi) \int_{\gamma-\omega_2 i}^{\gamma-\omega_1 i} s^{-1/2} \exp \left[ts - |\xi - x| s^{1/2}\right] ds d\xi \\
&\qquad + \int_0^{\infty} f(\xi) \int_{\gamma+\omega_1 i}^{\gamma+\omega_2 i} s^{-1/2} \exp \left[ts - |\xi - x| s^{1/2}\right] ds d\xi \Bigg|.
\end{aligned}
$$

Since the integrands are analytic in s, the integrations on s can be carried out over the contour $C_1+C_2+C_3+C_4+C_5+C_6$, where the C's are straight line segments with end points as follows: C_1: $\gamma-\omega_2 i$, $-\omega_2 i$; C_2: $-\omega_2 i$, $-\omega_1 i$; C_3: $-\omega_1 i$, $\gamma-\omega_1 i$; C_4: $\gamma+\omega_1 i$, $\omega_1 i$; C_5: $\omega_1 i$, $\omega_2 i$; C_6: $\omega_2 i$, $\gamma+\omega_2 i$.

The contribution of C_1 is

$$\int_0^{\infty} f(\xi) \int_{\gamma}^{0} (\zeta - \omega_2 i)^{-1/2} \exp \left[t\zeta - t\omega_2 i - |\xi - x| (\zeta - \omega_2 i)^{1/2}\right] d\zeta d\xi,$$

which is less in absolute value than

$$\int_0^{\infty} \alpha\Omega^{-1/2} \exp \left[\beta\xi + t\gamma - |\xi - x| (\Omega/2)^{1/2}\right]\gamma d\xi \leqq \frac{3\alpha(2)^{1/2} e^{t\gamma+\beta x}\gamma}{\Omega}.$$

A similar proof shows that the total contribution of $C_1+C_3+C_4+C_6$ is less than

$$\frac{12\alpha\gamma(2)^{1/2}e^{t\gamma+\beta x}}{\Omega},$$

which approaches zero as Ω approaches infinity, uniformly in x and t.

Finally, the contributions of C_3 and C_5 can be combined to give

$$2\int_0^\infty f(\xi)\int_{\omega_1}^{\omega_2} \exp\left[-\mid \xi - x\mid (\eta/2)^{1/2}\right] \cdot \cos\left[t\eta - (\pi/4) - \mid \xi - x\mid (\eta/2)^{1/2}\right]\eta^{-1/2}d\eta d\xi.$$

Split the range of integration at $\xi = \xi_1 > 2x_1$. The contribution of the integral from ξ_1 to ∞ is less in absolute value than

$$2\int_{\xi_1}^\infty \alpha e^{\beta\xi}\int_{\omega_1}^{\omega_2} \eta^{-1/2}\exp\left[-\xi(\eta/2)^{1/2}/2\right]d\eta d\xi.$$

On making the change of variable $\eta = 8\zeta^2\xi^{-2}$ this is found to be less than

$$8\alpha(2)^{1/2}\int_{\xi_1}^\infty \xi^{-1}\exp\left[\beta\xi - \xi(\Omega/2)^{1/2}/2\right]d\xi.$$

If $\Omega > 8\beta^2$ this integral converges, and hence can be made less than ϵ by a suitable choice of ξ_1. Choose such a ξ_1 and let it remain fixed.

We have remaining the integral from 0 to ξ_1:

$$\int_0^{\xi_1} f(\xi)\int_{\omega_1}^{\omega_2} \eta^{-1/2}\exp\left[-\mid \xi - x\mid (\eta/2)^{1/2}\right] \cdot\cos\left[t\eta - (\pi/4) - \mid \xi - x\mid (\eta/2)^{1/2}\right]d\eta d\xi.$$

On making the substitution

$$(t\eta)^{1/2} - \mid \xi - x\mid /2(2t)^{1/2} = \zeta$$

and using the addition formula for the cosine, this is seen to be less in absolute value than

$$2t^{-1/2}\left|\int_0^{\xi_1} f(\xi)\exp\left[-(\xi - x)^2/4t\right]\cos\left[(\pi/4) + (\xi - x)^2/8t\right] \cdot\int_{\zeta_1}^{\zeta_2}\exp\left[-\zeta\mid \xi - x\mid /(2t)^{1/2}\right]\cos \zeta^2 d\zeta d\xi\right|$$

plus the same expression with cos replaced by sin, where

$$\zeta_n = (t\omega_n)^{1/2} - \mid \xi - x\mid /2(2t)^{1/2}, \qquad n = 1, 2.$$

By the second mean-value theorem this is equal to

$$
(19)\qquad \begin{aligned} 2t^{-1/2}\Bigg| \int_0^{\xi_1} f(\xi) \exp\,[-\,|\,\xi - x\,|\,(\omega_1/2)^{1/2}] \cos\,[(\pi/4) + (\xi - x)^2/8t] \\ \cdot \int_{\zeta_1}^{\zeta_3} \cos \zeta^2 d\zeta d\xi \Bigg| \end{aligned}
$$

plus the same expression with cos replaced by sin, and $\zeta_1 \leqq \zeta_3 \leqq \zeta_2$.

But the integrals $\int_0^\infty \sin \zeta^2 d\zeta$ and $\int_0^\infty \cos \zeta^2 d\zeta$ converge; hence there is a positive number M such that

$$
\left| \int_{\zeta_1}^{\zeta_3} \cos \zeta^2 d\zeta \right| < M > \left| \int_{\zeta_1}^{\zeta_3} \sin \zeta^2 d\zeta \right|
$$

for any values of ζ_1, ζ_3. Thus the expression (19) is less than

$$
4Mt^{-1/2} \int_0^{\xi_1} |\,f(\xi)\,| \exp\,[-\,|\,\xi - x\,|\,(\omega_1/2)^{1/2}]d\xi.
$$

On splitting the range of integration at $\xi = x$ this is seen to be less than

$$
12(2)^{1/2} M e^{bx}/(t\Omega)^{1/2},
$$

which approaches zero as Ω approaches infinity, uniformly in x and t, q.e.d.

LEMMA 3. *If $0 \leqq x \leqq x_1$, $0 < t_1 \leqq t \leqq t_2$, then $L^{-1}\{\partial g(x, s)/\partial x\}$ exists and converges uniformly in x and t.*

Proof. Using the method and notation of the proof of Lemma 2, it is easily found that the contribution of $C_1 + C_3 + C_4 + C_6$ is less in absolute value than

$$
12\alpha\gamma e^{t\gamma + \beta x}(2/\Omega)^{1/2}.
$$

Similarly, the contribution of $C_2 + C_5$ over the range $\xi_1 \leqq \xi \leqq \infty$ is less than

$$
96\alpha\xi_1^{-1} \exp\,[-\,\xi_1(\Omega/2)^{1/2}/4].
$$

Choose a fixed ξ_1 such that this is less than ϵ. Then we have remaining the expression

$$
\begin{aligned} 2t^{-1} \int_0^{\xi_1} |\,f(\xi)\,| \cdot \Bigg| \int_{\zeta_1}^{\zeta_2} \exp\,[-\,|\,\xi - x\,|\,\{\zeta + |\,\xi - x\,|\,/2(2t)^{1/2}\}/(2t)^{1/2}] \\ \cdot \cos\,[\zeta^2 - (\xi - x)^2/8t] \cdot [\zeta + |\,\xi - x\,|\,/2(2t)^{1/2}]d\zeta \Bigg| d\xi. \end{aligned}
$$

By means of the last bracket this can be separated into two terms, the second of which, as in the proof of Lemma 2, is less than

$$
12\alpha M \xi_1 e^{\beta x}(2/t^3\Omega)^{1/2}.
$$

This leaves:

$$2\alpha t^{-1}\int_0^{\xi_1} e^{\beta\xi}\left|\int_{\zeta_1}^{\zeta_2}\exp\left[-|\xi-x|\left\{\zeta+|\xi-x|/2(2t)^{1/2}\right\}/(2t)^{1/2}\right]\right.$$

$$\left.\cdot\cos\left[\zeta^2-(\xi-x)^2/8t\right]\zeta d\zeta\right|d\xi.$$

Integrating by parts with respect to ζ gives

$$\alpha t^{-1}\int_0^{\xi_1} e^{\beta\xi}\left\{\exp\left[-|\xi-x|(\omega_2/2)^{1/2}\right]+\exp\left[-|\xi-x|(\omega_1/2)^{1/2}\right]\right\}d\xi$$

$$+\alpha(2t^3)^{-1/2}\int_0^{\xi_1} e^{\beta\xi}\left|\int_{\zeta_1}^{\zeta_2}|\xi-x|\exp\left[-|\xi-x|\left\{\zeta+|\xi-x|/2(2t)^{1/2}\right\}/(2t)^{1/2}\right]\right.$$

$$\left.\cdot\sin\left[\zeta^2-(\xi-x)^2/8t\right]d\zeta\right|d\xi,$$

which, as before, is less than

$$6\alpha e^{\beta x}t^{-1}(2/\Omega)^{1/2}+6\alpha M\xi_1 e^{\beta x}/(t^3\Omega)^{1/2},$$

which approaches zero as Ω approaches infinity, uniformly in x and t, q.e.d.

The following can be proved by the same methods:

LEMMA 4. *If* $0\leqq x\leqq x_1$, $0<t_1\leqq t\leqq t_2$, $n=0, 1$, *then* $L^{-1}\{\partial^n z(x, s)/\partial x^n\}$ *exists and converges uniformly in* x *and* t.

Since $|sg(x, s)|$ does not approach zero as $|s|$ approaches infinity, the Laplace transformation method cannot be used to establish the existence(6) of $\partial G/\partial t$ and $\partial^2 G/\partial x^2$. This will be done by classical methods by means of the following:

LEMMA 5. $g(x, s)=L\{G(x, t)\}$, *where*

$$G(x, t)=\int_0^\infty f(\xi)\Theta(x-\xi, t)d\xi,$$

and then $G(x, t)=L^{-1}\{g(x, s)\}$.

Proof. The substitution $\xi=x+2\zeta t^{1/2}$ shows that

$$\int_0^\infty |f(\xi)\Theta(x-\xi, t)|\,d\xi\leqq 2\alpha\exp\left[\beta x+\beta^2 t\right]. \tag{20}$$

Hence $L\{G\}$ exists and converges absolutely. By (20) the order of integration in $L\{G\}$ can be inverted, giving(7)

---

(6) Cf. [C, Theorem 6].
(7) Cf. [D, Appendix 2, table].

$$L\{G\} = \int_0^\infty f(\xi)L\{\Theta(x-\xi,t)\}d\xi = g(x,s).$$

Hence, by [D, chap. 6, §5, Theorem 2], G is given by the inversion integral $L^{-1}\{g\}$, q.e.d.

Using these lemmas we are now in a position to establish the solution of the original boundary value problem.

5. **The solution established.**

THEOREM 1. *The function $W_\nu(x,t) \equiv L^{-1}\{w_\nu(x,s)\}$ exists and has the following properties*:

(a) $W_1(x,t) = (1/2)\int_{-\infty}^0 f(\xi)\{\Theta(x-\xi, a_1t)+\Theta(x+\xi, a_1t)\}d\xi$ *and W_2 is obtained from W_1 upon replacing a_1 by a_2 and integrating from* 0 *to* $+\infty$.

(b) *$W_\nu(x,t)$ satisfies the differential equation* (1).

(c) *$W_\nu(x,t)$ satisfies the initial condition* (2) *at any point of continuity of $f(x)$*.

(d) $\lim_{x\to 0} \partial W_\nu/\partial x = 0$.

Proof. Equation (a) follows from Lemma 5, and the remainder of the theorem is then easily proved by classical methods([8]).

To obtain the complete solution of the problem it is now sufficient to prove the following:

THEOREM 2. *The function $Y_\nu(x,t) \equiv L^{-1}\{y_\nu(x,s)\}$ exists and has the following properties*:

(a) *$Y_\nu(x,t)$ satisfies the differential equation* (1).

(b) $\lim_{t\to 0} Y_\nu(x,t) = 0, \qquad x \neq 0.$

(c) $\lim_{x\to -0} k_1 \dfrac{\partial Y_1}{\partial x} = \lim_{x\to +0} k_2 \dfrac{\partial Y_2}{\partial x}, \qquad t > 0.$

(d) $Y_2(+0,t) + W_2(+0,t) - Y_1(-0,t) - W_1(-0,t) = \lim_{x\to 0} \lambda k_\nu \dfrac{\partial Y_\nu}{\partial x}, \qquad t > 0.$

Proof. Y_ν exists by Lemma 1.

(a) By Lemma 1 and inequality (16), Y_ν satisfies the hypotheses of [C, Theorems 6 and 9]. Hence, if $x\neq 0$, $t>0$, $\partial Y_\nu/\partial t = L^{-1}\{sy_\nu\}$. But $\partial^2 Y_\nu/\partial x^2 = L^{-1}\{\partial^2 y_\nu/\partial x^2\}$, and, since (11) can be differentiated inside the integral sign,

$$\frac{\partial^2 y_\nu}{\partial x^2} = a_\nu^{-1} s y_\nu, \text{ q.e.d.}$$

(b) By Lemma 1 and inequality (16) the hypotheses of [C, Theorem 4] are satisfied, from which the result follows immediately.

([8]) Cf. Carslaw, loc. cit., §§16–18 for a proof under more stringent conditions on $f(x)$.

(c) and (d) From equations (11) and (13) it is easily seen that

$$\lim_{x\to -0} k_1 \frac{\partial y_1(x, s)}{\partial x} = \lim_{x\to +0} k_2 \frac{\partial y_2(x, s)}{\partial x}, \tag{21}$$

$$y_2(+0, s) + w_2(+0, s) - y_1(-0, s) - w_1(-0, s) = \lim_{x\to 0} \lambda k_\nu \frac{\partial y_\nu}{\partial x}. \tag{22}$$

By Lemma (2) and inequality (17) $w_\nu(x, s)$ satisfies the hypotheses of [C, Theorem 8]. Hence $W_\nu(\pm 0, t) = L^{-1}\{w_\nu(\pm 0, s)\}$, $t>0$.

By Lemma 4 and inequality (16), $y_\nu(x, s)$ and $\partial y_\nu/\partial x$ satisfy the hypotheses(9) of [C, Theorems 8 and 9]. Hence $Y_\nu(\pm 0, t) = L^{-1}\{y_\nu(\pm 0, s)\}$, $t>0$ and

$$\lim_{x\to\pm 0} \frac{\partial Y_\nu}{\partial x} = L^{-1}\left\{\lim_{x\to\pm 0} \frac{\partial y_\nu}{\partial x}\right\}, \qquad t > 0.$$

Hence conditions (c) and (d) follow from (21) and (22), q.e.d.

To summarize, the solution of the boundary value problem (1)–(4) is given by

$$U(x, t) = W_\nu(x, t) + Y_\nu(x, t) \tag{10'}$$

where

$$W_1(x, t) = \frac{1}{2}\int_{-\infty}^{0} f(\xi)\{\Theta(x - \xi, a_1 t) + \Theta(x + \xi, a_1 t)\}d\xi, \tag{13'}$$

$W_2(x, t)$ is obtained from $W_1(x, t)$ upon replacing a_1 by a_2 and integrating from 0 to $+\infty$, Θ is given by (9), and

$$Y_\nu(x, t) = \lim_{\omega\to\infty} \frac{1}{2\pi i}\int_{\gamma-\omega i}^{\gamma+\omega i} e^{ts} y_\nu(x, s)ds, \tag{23}$$

if $\gamma > 32\beta^2$, where $y_\nu(x, s)$ is given by equations (11).

6. **Explicit forms of** Y. In the preceding section the function $Y_\nu(x, t)$ was obtained as a complex inversion integral. We now develop two more practical formulas.

THEOREM 3. *$Y_\nu(x, t)$ is given by the following formulas:*

$$\begin{aligned} Y_1(x, t) = k_2 a_2^{-1/2} \int_0^\infty e^{-A\mu}\Big\{a_2^{-1/2} \int_0^\infty f(\xi)\Theta(\,|\,\xi a_2^{-1/2} - x a_1^{-1/2}\,| + B\mu, t)d\xi \\ - a_1^{-1/2}\int_{-\infty}^{0} f(\xi)\Theta(B\mu + |\,x + \xi\,|\, a_1^{-1/2}, t)d\xi\Big\}d\mu, \qquad x < 0 < t, \end{aligned} \tag{24}$$

(9) Actually, in [C] it is assumed that $\partial y(x, s)/\partial x = O(|s|^{-1})$, but the weaker condition $\partial y/\partial x = O(|s|^{-1/2})$ given by our inequality (16), is sufficient, as can be seen from an examination of Churchill's proof.

and Y_2 is obtained from Y_1 upon interchanging the subscripts 1 *and* 2 *and the integrals* $\int_0^\infty d\xi$ *and* $\int_{-\infty}^0 d\xi$.

Proof. Note that in equations (11) each term is of the form

$$j(x, s) \equiv s^{-1/2}(A + Bs^{1/2})^{-1} \int_0^\infty f(\xi) \exp\,[-(x+\xi)s^{1/2}]d\xi, \qquad x > 0.$$

Since $(A+Bs^{1/2})^{-1} = \int_0^\infty \exp\,[-(A+Bs^{1/2})\mu]d\mu$, we can write

$$j(x, s) = \int_0^\infty e^{-A\mu} \int_0^\infty f(\xi)s^{-1/2} \exp\,[-(x+\xi+B\mu)s^{1/2}]d\xi d\mu. \tag{25}$$

By Lemma 1, $J(x, t) = L^{-1}\{j(x, s)\}$ exists. Hence, to prove the present theorem it is sufficient to show that the integration on s can be performed inside the inner integral in (25). This will be done in two steps.

First we show that the integrations on μ and s can be inverted. It is sufficient to prove that[19]

$$\lim_{\mu_1 \to \infty} \int_0^{\mu_1} e^{-A\mu}e^{ts} \int_0^\infty s^{-1/2}f(\xi) \exp\,[-(x+\xi+B\mu)s^{1/2}]d\xi d\mu \tag{26}$$

converges uniformly in s, and

$$\lim_{\omega \to \infty} \int_{\gamma-\omega i}^{\gamma+\omega i} e^{-A\mu}e^{ts} \int_0^\infty s^{-1/2}f(\xi) \exp\,[-(x+\xi+B\mu)s^{1/2}]d\xi ds \tag{27}$$

converges uniformly in μ.

To prove (26) note that the absolute value of the integral from μ_1 to ∞ is less than

$$\int_{\mu_1}^\infty e^{-A\mu}e^{t\gamma}2^{3/2}\alpha\,|\,s\,|^{-1} \exp\,[-(x+B\mu)(|\,s\,|/2)^{1/2}]d\mu \tag{28}$$

by inequality (16), with $n = 1/2$ and x replaced by $x+B\mu$. But (28) is less than

$$\frac{2^{3/2}\alpha e^{t\gamma}|\,s\,|^{-1} \exp\,[-x(|\,s\,|/2)^{1/2}] \exp\,[-\{A+B(|\,s\,|/2)^{1/2}\}\mu_1]}{A+B(|\,s\,|/2)^{1/2}},$$

which obviously approaches zero as μ_1 approaches infinity, uniformly in s.

To prove (27), the absolute value of the double integral is less than

$$\int_{\gamma-\infty i}^{\gamma+\infty i} 2^{1/2}\alpha e^{t\gamma}|\,s\,|^{-1} \exp\,[-x(|\,s\,|/2)^{1/2}]ds$$

by inequality (16). But this converges and is independent of μ, q.e.d. Hence

---

[19] Cf. Pierpont, *Theory of functions of real variables*, 1905, vol. 1 p. 489 §680.3.

$$J(x,t) = \int_0^\infty e^{-A\mu} \lim_{\omega\to\infty} \frac{1}{2\pi i}\int_{\gamma-\omega i}^{\gamma+\omega i} e^{ts} s^{-1/2}\int_0^\infty f(\xi)\exp\left[-(x+\xi+B\mu)s^{1/2}\right]d\xi ds d\mu.$$

Now, by a proof similar to that of Lemma 5, or by inequalities (16) and (18), the inversion of the integrations on ξ and s can be justified. Therefore

$$J(x,t) = \int_0^\infty e^{-A\mu}\int_0^\infty f(\xi)L^{-1}\{s^{-1/2}\exp\left[-(x+\xi+B\mu)s^{1/2}\right]\}d\xi d\mu.$$

But the inner integral is well known[11]. Hence,

$$J(x,t) = \int_0^\infty e^{-A\mu}\int_0^\infty f(\xi)\Theta(x+\xi+B\mu, t)d\xi d\mu, \text{ q.e.d.}$$

Another form of Y is given by the following

THEOREM 4.

$$\begin{aligned} Y_1(x,t) = (k_2/Ba_2^{1/2})\int_0^t &\exp\left[A^2(t-\tau)B^{-2}\right] \\ &\cdot\Big\{a_2^{-1/2}\int_0^\infty f(\xi)\Psi(\xi a_2^{-1/2} - x a_1^{-1/2}, \tau)d\xi \\ &+ a_1^{-1/2}\int_{-\infty}^0 f(\xi)\Psi([x+\xi]a_1^{-1/2},\tau)d\xi \\ &- (A/Ba_2^{1/2})\int_0^\infty f(\xi)\Theta(\xi a_2^{-1/2} - x a_1^{-1/2}, \tau)d\xi \\ &+ AB^{-1}\int_{-\infty}^0 f(\xi)\Theta(x+\xi, a_1\tau)d\xi\Big\}d\tau, \qquad x<0<t, \end{aligned} \tag{29}$$

and $Y_2(x,t)$ is obtained from $Y_1(x,t)$ on interchanging subscripts 1 and 2 and the integrals $\int_0^\infty$ and $\int_{-\infty}^0$.

Proof. On observing that $(A+Bs^{1/2})^{-1} = (s-A^2B^{-2})^{-1}[s^{1/2}B^{-1} - AB^{-2}]$, and that $L^{-1}\{(s-A^2B^{-2})^{-1}\} = \exp\left[A^2tB^{-2}\right]$, the result follows immediately from the "Faltung theorem[12]" since all the Laplace transformations concerned are absolutely convergent. (Compare the proof of Lemma 5 above.)

UNIVERSITY OF CALIFORNIA,
DAVIS, CALIF.

(11) Cf. [D, Appendix 2, table].
(12) Cf. [D, chap. 8. §5, Theorem IV$_b$].

CONGRUENCES IN UNITARY SPACE

BY

N. COBURN

1. **Introduction.** In this paper, we shall study the properties of a congruence of ∞^{n-1} curves which are imbedded in a unitary space of n dimensions K_n (a real topological space of $2n$ dimensions). First, we consider the general case—when the curves are ∞^{n-1} unitary curves K_1 (real topological spaces of two dimensions)—and determine the associated congruence affinors. Then, we determine the necessary and sufficient conditions in terms of congruence vectors that the ∞^{n-1} congruence curves should be either unitary U_1 (unitary Euclidean curves)(1) or real curves X_1 (real topological spaces of one dimension). If the curves of the congruence are all real X_1, then we define the congruence to be real; if the curves are all unitary U_1, then we define the congruence to be complex Euclidean.

In the next section, we study two systems of Pfaffians which enable us to define two types of orthogonality: (1) ∞^1 hypersurfaces which are *completely unitary orthogonal* to the congruence curves; (2) ∞^1 hypersurfaces which are *semi-unitary orthogonal* to the congruence curves. It is shown that: (1) the ∞^1 hypersurfaces which are completely unitary orthogonal to the congruence curves admit of an intrinsic parameterization and are ∞^1 unitary K_{n-1}; (2) if the ∞^1 hypersurfaces which are semi-unitary orthogonal to the congruence curves admit of a parameterization, then they constitute ∞^1 semi-analytic(2) spaces X_{n-1}. A further analytical characterization of these two types of surfaces is given.

The remainder of our work deals with two problems: (1) a characterization in terms of congruence affinors of those congruences which are either completely unitary orthogonal or semi-unitary orthogonal to ∞^1 hypersurfaces in K_n; (2) special properties of these two types of congruences. Thus, in connection with the second problem, it is shown that if the congruence is either real, or complex Euclidean, analytic and completely unitary orthogonal to ∞^1 hypersurfaces, then the conditions satisfied by the congruence vector are similar to those satisfied by the congruence vector which is orthogonal to ∞^1 hypersurfaces(3) in V_n (n-dimensional Riemannian space). Again, if: (1) the congruence is real and geodesic; (2) the K_n has a symmetric connection, then every two hypersurfaces which are semi-unitary orthogonal to the

Presented to the Society, February 28, 1942; received by the editors February 3, 1942.

(1) [5, vol. 2, p. 251].

(2) [3, equation (2.10)].

(3) [5, vol. 2, p. 28, equation 5.2].

congruence intercept equal arc segments on all X_1 of the congruence. This latter result is similar to a theorem(4) in Riemannian space.

2. **Notation**(5). Consider a real space of $2n$ dimensions X_{2n} whose coordinates are given by the real variables

$$(2.1)\qquad x^\lambda,\ y^\lambda, \qquad \lambda, \mu = 1, 2, \cdots, n.$$

Into this X_{2n}, we introduce the complex coordinates

$$(2.2)\qquad \xi^\lambda = x^\lambda + iy^\lambda, \qquad i = (-1)^{1/2},$$

$$(2.3)\qquad \xi^{\lambda^*} = x^\lambda - iy^\lambda.$$

Since the Jacobian of this transformation $(-2i)$ does not vanish over X_{2n}, the ξ^λ, ξ^{λ^*} constitute a set of $2n$ independent variables which map the X_{2n}. In view of the fact that ξ^{λ^*} are complex conjugate to ξ^λ, we can determine the points of X_{2n} by assigning complex numbers to merely ξ^λ. Hence, we say that the ξ^λ determine "points" which build a complex space of n dimensions (the above real topological X_{2n}).

Let us denote partial derivatives by

$$(2.4)\qquad \partial_\mu = \partial/\partial\xi^\mu, \qquad \partial_{\mu^*} = \partial/\partial\xi^{\mu^*}.$$

If $\psi(\xi^\lambda, \xi^{\lambda^*})$ is an analytic function of ξ^λ, ξ^{λ^*}, then we shall say that ψ is semianalytic; if $\phi(\xi^\lambda)$ is an analytic function of ξ^λ (or ξ^{λ^*}) alone, then we shall say that ϕ is analytic. In view of (2.4), we may express this last condition by

$$(2.5)\qquad \partial_{\mu^*}\phi = 0.$$

One further important formal idea must be noted—that of the conjugate function and equation. If we replace i by $-i$ in ϕ, the resulting function is denoted by ϕ^* (where ϕ is a scalar). From (2.2), (2.3), we see that ξ^λ must be replaced by ξ^{λ^*} and vice versa. Hence $\phi(\xi^\lambda)$ becomes $\phi^*(\xi^{\lambda^*})$. In the case of affinors, the conjugate affinor is obtained in the same manner. However, we shall indicate this conjugate by starring the previously unstarred indices and removing the star from the previously starred indices. Thus the conjugate of $v_{\lambda\mu^*}$ is $v_{\lambda^*\mu}$. Furthermore from our discussion, it follows that

$$(2.6)\qquad \partial_\mu\phi^* = 0.$$

The equation (2.6) is the so-called conjugate equation to (2.5). Also to every affinor equation, there corresponds a conjugate equation obtained by replacing i by $-i$ and hence each affinor by its conjugate. The truth of this last statement can be seen by decomposing each affinor into its real and imaginary parts(6). In the following, we shall indicate the validity of the conjugate equation by the abbreviation "conj."

(4) [5, vol. 2, p. 31].

(5) Our notation is that of [5].

(6) Note, by composite differentiation, it follows that $\partial_\mu = \partial/\partial x^\mu - i\partial/\partial y^\mu$, $\partial_{\mu^*} = \partial/\partial x^\mu + i\partial/\partial y^\mu$.

We specify that the group of this complex X_n shall be the analytic group(7) of coordinate transformations. Now, let us introduce a connection in X_n by means of the n^3 quantities $\Gamma^{\lambda}_{\mu\alpha}$ which are functions of the coordinates ξ^{λ}, ξ^{λ^*}. We define the covariant differential of a contravariant vector $v^{\lambda}(\xi^{\lambda}, \xi^{\lambda^*})$ by

$$(2.7)\qquad \delta v^{\lambda} = dv^{\lambda} + \Gamma^{\lambda}_{\mu\alpha} v^{\alpha} d\xi^{\mu}, \text{ conj.}$$

Likewise, we define the covariant differential of a covariant vector $w_{\lambda}(\xi^{\lambda}, \xi^{\lambda^*})$ by

$$(2.8)\qquad \delta w_{\lambda} = dw_{\lambda} - \Gamma^{\alpha}_{\mu\lambda} w_{\alpha} d\xi^{\mu}, \text{ conj.}$$

By expanding the ordinary differential of a vector, we obtain

$$(2.9)\qquad dv^{\lambda} = d\xi^{\mu}\partial_{\mu}v^{\lambda} + d\xi^{\mu^*}\partial_{\mu^*}v^{\lambda}, \text{ conj.},$$

$$(2.10)\qquad dw_{\lambda} = d\xi^{\mu}\partial_{\mu}w_{\lambda} + d\xi^{\mu^*}\partial_{\mu^*}w_{\lambda}, \text{ conj.}$$

If we define the covariant derivative of v^{λ}, w_{λ} by the equations

$$(2.11)\qquad \nabla_{\mu}v^{\lambda} = \partial_{\mu}v^{\lambda} + \Gamma^{\lambda}_{\mu\alpha}v^{\alpha}, \qquad \nabla_{\mu}w_{\lambda} = \partial_{\mu}w_{\lambda} - \Gamma^{\alpha}_{\mu\lambda}w_{\alpha}, \text{ conj.},$$

$$(2.12)\qquad \nabla_{\mu^*}v^{\lambda} = \partial_{\mu^*}v^{\lambda}, \qquad \nabla_{\mu^*}w_{\lambda} = \partial_{\mu^*}w_{\lambda}, \text{ conj.},$$

then by use of the equations (2.9) through (2.12), the equations (2.7), (2.8) become

$$(2.13)\qquad \delta v^{\lambda} = d\xi^{\mu}\nabla_{\mu}v^{\lambda} + d\xi^{\mu^*}\nabla_{\mu^*}v^{\lambda}, \text{ conj.},$$

$$(2.14)\qquad \delta w_{\lambda} = d\xi^{\mu}\nabla_{\mu}w_{\lambda} + d\xi^{\mu^*}\nabla_{\mu^*}w_{\lambda}, \text{ conj.}$$

An hermitian X_n with covariant derivative defined by (2.11), (2.12) is denoted by K_n.

Let us introduce an hermitian tensor with hermitian symmetry, that is,

$$(2.15)\qquad a_{\lambda\mu^*} = [(a_{\lambda\mu^*})^*]' = a_{\mu^*\lambda},$$

the sign ($'$) indicating the transpose matrix. If we condition the $a_{\lambda\mu^*}$ by requiring that

$$(2.16)\qquad \delta a_{\lambda\mu^*} = 0 = (\partial_{\nu}a_{\lambda\mu^*} - \Gamma^{\rho}_{\nu\lambda}a_{\rho\mu^*})d\xi^{\nu} + (\partial_{\nu^*}a_{\lambda\mu^*} - \Gamma^{\rho^*}_{\nu^*\mu^*}a_{\lambda\rho^*})d\xi^{\nu^*},$$

then the space K_n is said to be a unitary K_n. For such a space from (2.16), we can show(8)

$$(2.17)\qquad \nabla_{\nu}a_{\lambda\mu^*} = \partial_{\nu}a_{\lambda\mu^*} - \Gamma^{\rho}_{\nu\lambda}a_{\rho\mu^*} = 0,$$

$$(2.18)\qquad \nabla_{\nu^*}a_{\lambda\mu^*} = \partial_{\nu^*}a_{\lambda\mu^*} - \Gamma^{\rho^*}_{\nu^*\mu^*}a_{\lambda\rho^*} = 0.$$

The $a_{\lambda\mu^*}$ is now a fundamental tensor and can be used to raise and lower in-

(7) The analytic group of transformations is given by $\xi^{\lambda'} = \xi^{\lambda'}(\xi^{\alpha})$, conj.

(8) [5, vol. 2, p. 234].

dices through the ∇ operator. If we define the contravariant fundamental tensor $a^{\nu^*\lambda}$ by

$$(2.19)\qquad a^{\nu^*\lambda} a_{\mu\nu^*} = A_\mu^\lambda, \text{ conj.},$$

where A_μ^λ is the unit affinor, then (2.17), (2.18) may be solved for the connection

$$(2.20)\qquad \Gamma_{\mu\lambda}^{\kappa} = a^{\nu^*\kappa} \partial_\mu a_{\lambda\nu^*},$$

$$(2.21)\qquad \Gamma_{\mu^*\lambda^*}^{\kappa^*} = a^{\kappa^*\nu} \partial_{\mu^*} a_{\nu\lambda^*}.$$

Finally, we introduce the torsion affinor

$$(2.22)\qquad S_{\mu\lambda}^{\cdot\cdot\kappa} = (1/2)(\Gamma_{\mu\lambda}^{\kappa} - \Gamma_{\lambda\mu}^{\kappa}) = \Gamma_{[\mu\lambda]}^{\kappa}, \text{ conj.}$$

The sign [ ] means that the antisymmetric product of the enclosed indices is to be formed; the sign | | enclosing indices means that those indices are to be excluded in forming the antisymmetric product. When the torsion affinor can be written as

$$(2.23)\qquad S_{\mu\lambda}^{\cdot\cdot\kappa} = A_{[\mu}^{\kappa} p_{\lambda]}, \text{ conj.},$$

the unitary space K_n is said to have a semi-symmetric connection.

3. **Congruences in unitary K_n.** Consider a vector field $u^\lambda(\xi^\lambda, \xi^{\lambda^*})$ defined over the unitary K_n. The system of differential equations in the parameter t,

$$(3.1)\qquad d\xi^\lambda/u^\lambda = dt, \text{ conj.},$$

is said to define a congruence(9) of ∞^{n-1} curves in the unitary K_n. We shall study the decomposition of the affinors $\nabla_\alpha u_\lambda$, $\nabla_{\alpha^*} u_\lambda$. Consider affinors $l_{\alpha\lambda}$, $l_{\alpha^*\lambda}$ which we define as the projections of $\nabla_\alpha u_\lambda$, $\nabla_{\alpha^*} u_\lambda$, respectively, upon the local U_{n-1} which is unitary orthogonal(10) to u_λ. Hence, it follows that

$$(3.2)\qquad u^\alpha l_{\alpha\lambda} = 0, \qquad u^{\alpha^*} l_{\alpha^*\lambda} = 0, \text{ conj.},$$

$$(3.3)\qquad u^\mu l_{\alpha\mu} = 0, \qquad u^\mu l_{\alpha^*\mu} = 0, \text{ conj.}$$

Furthermore, let w_α, z_α, x_α, y_α be four arbitrary vectors in the above local U_{n-1}, that is,

$$(3.4)\qquad w_\alpha u^\alpha = z_\alpha u^\alpha = x_\alpha u^\alpha = y_\alpha u^\alpha = 0, \text{ conj.}$$

We can now write(11)

$$(3.5)\qquad \nabla_\alpha u_\lambda = l_{\alpha\lambda} + u_\alpha w_\lambda + z_\alpha u_\lambda + p u_\alpha u_\lambda, \text{ conj.},$$

(9) [5, vol. 2, p. 27, equation 5.1].

(10) This local U_{n-1} is determined by those vectors u_j^λ (subscript $j=1, \cdots, n-1$) which are solutions of $u_j^\lambda u_\lambda = 0$.

(11) [5, vol. 1, p. 19, §k].

$$\nabla_{\alpha^*} u_\lambda = l_{\alpha^*\lambda} + u_{\alpha^*} x_\lambda + y_{\alpha^*} u_\lambda + q^* u_{\alpha^*} u_\lambda, \text{ conj.,} \tag{3.6}$$

where p, q are scalars.

If the parameter t in (3.1) is complex and the congruence curves are U_1 or if the parameter t is real and hence the curves are X_1, then an analytic arc length parameter s exists(12)

$$s = s(t), \text{ conj.} \tag{3.7}$$

Now, if we replace the parameter t by s in the ∞^{n-1} congruence curves, then the associated congruence vector u^λ (we indicate the vector by the same symbol as before) is a unit vector, that is,

$$u^\lambda u_\lambda = 1, \text{ conj.} \tag{3.8}$$

Because of (3.8), certain relations exist between the affinors in (3.5), (3.6). Before finding these relations, we formulate

DEFINITION 1. (a) *If the parameter t in* (3.1) *is real, then the congruence defined by* (3.1) *will be said to be real. This congruence consists of $\infty^{n-1}X_1$ in K_n*; (b) *if the parameter t is complex but the ∞^{n-1} curves of the congruence are U_1, then we shall say that the congruence is complex Euclidean.*

By covariant differentiation of (3.8), we obtain

$$(\nabla_\alpha u_\lambda)u^\lambda = -(\nabla_\alpha u^\lambda)u_\lambda, \text{ conj.} \tag{3.9}$$

As a consequence of the equation

$$u^\lambda = a^{\beta^*\lambda} u_{\beta^*}, \text{ conj.,} \tag{3.10}$$

we find that the right-hand side of (3.9) can be expressed in terms of $\nabla_\alpha u_{\beta^*}$, that is,

$$(\nabla_\alpha u_\lambda)u^\lambda = -(\nabla_\alpha u_{\beta^*})u^{\beta^*}, \text{ conj.} \tag{3.11}$$

By use of (3.5), (3.6), the relation (3.11) can be shown to be equivalent to

$$z_\alpha = -y_\alpha, \text{ conj.,} \tag{3.12}$$

$$p = -q, \text{ conj.} \tag{3.13}$$

Conversely, if (3.12), (3.13) are valid, then the validity of (3.11), (3.9) follows. But (3.9) may be written in the form

$$\nabla_\alpha(u_\lambda u^\lambda) = 0, \text{ conj.} \tag{3.14}$$

Hence, it follows that

$$u_\lambda u^\lambda = c, \text{ conj.,} \tag{3.15}$$

where c is some arbitrary constant in the unitary K_n. By use of (3.1), the

(12) [1, Theorems 3, 4].

equation (3.15) becomes

$$a_{\lambda_\mu \cdot} d_t \xi^\lambda d_{t^*} \xi^{\mu^*} = c. \tag{3.16}$$

But this means that the curves of the congruence are ∞^{n-1} U_1 (for complex t)(12) or ∞^{n-1} X_1 (for real t). Hence, we have the theorem.

THEOREM 1. *The necessary and sufficient conditions that the solutions u_λ of (3.5), (3.6)—when they exist—should define either a real congruence or a complex Euclidean congruence is that $z_\alpha = -y_\alpha$, $p = -q$.*

4. Two systems of Pfaffians. Let us consider a general congruence vector $u_\lambda(\xi^\lambda, \xi^{\lambda^*})$. In the first place, we associate with this vector a system of two Pfaffians

$$u_\lambda d\xi^\lambda = 0, \tag{4.1}$$

$$u_{\lambda^*} d\xi^{\lambda^*} = 0. \tag{4.2}$$

Assuming that u_1, u_{1^*} do not vanish over some region D of the unitary K_n, we can rewrite the two previous equations in the form

$$d\xi^1 = -\sum{}'(u_\alpha/u_1)d\xi^\alpha, \qquad \alpha = 2, \cdots, n, \tag{4.3}$$

$$d\xi^{1^*} = -\sum{}'(u_{\alpha^*}/u_{1^*})d\xi^{\alpha^*}, \tag{4.4}$$

where $\sum'$ denotes summation over all repeated indices with the exception of the index 1. If the integrability conditions of this system are satisfied, we can solve(13) for ξ^1, ξ^{1^*}

$$\xi^1 = \xi^1(\xi^\alpha, \xi^{\alpha^*}, \underset{0}{\xi^1}, \underset{0}{\xi^{1^*}}), \qquad \alpha = 2, \cdots, n, \tag{4.5}$$

$$\xi^{1^*} = \xi^{1^*}(\xi^\alpha, \xi^{\alpha^*}, \underset{0}{\xi^1}, \underset{0}{\xi^{1^*}}), \tag{4.6}$$

where: (1) ξ^1, ξ^{1^*} (subindex 0) are arbitrary constants; (2) ξ^α, ξ^{α^*} ($\alpha = 2, \cdots, n$) are the independent variables; (3) ξ^1, ξ^{1^*} are the dependent variables. By solving for ξ^1, ξ^{1^*} (subindex 0), we obtain the two independent integrals of (4.1), (4.2),

$$\underset{0}{\xi^1} = f(\xi^\lambda, \xi^{\lambda^*}), \qquad \lambda = 1, \cdots, n, \tag{4.7}$$

$$\underset{0}{\xi^{1^*}} = g(\xi^\lambda, \xi^{\lambda^*}). \tag{4.8}$$

We now prove

LEMMA 1. *The two independent integrals f, g are conjugate functions.*

The equations (4.5), (4.6) become identities when the variables ξ^λ, ξ^{λ^*} ($\lambda = 1, 2, \cdots, n$) are replaced by the arbitrarily assigned constants ξ^λ, ξ^{λ^*}

(12) [4, p. 49].

(subindex 0). Hence (4.7), (4.8) become identities when the same substitution is made. That is, the functions f and g reduce to conjugate quantities ξ^1, ξ^{1*} (subindex 0) when the ξ^λ, $\xi^{\lambda*}$ $(\lambda=1, 2, \cdots, n)$ are assigned arbitrary values. Thus, f and g are conjugate functions.

Since the quantity ξ^{1*} (subindex 0) is known when the quantity ξ^1 (subindex 0) has been assigned some arbitrary value, we shall say that (4.7), (4.8) determine ∞^1 integrals of (4.1), (4.2). We prove

LEMMA 2. *The ∞^1 integrals of* (4.1), (4.2) *determine ∞^1 unitary hypersurfaces K_{n-1} in K_n.*

Let us denote ξ^α, $\xi^{\alpha*}$ $(\alpha=2, \cdots, n)$ by u^a, u^{a*} $(a=1, \cdots, n-1)$, respectively; these u^a, u^{a*} will serve as the hypersurface intrinsic parameters. Thus, (4.7), (4.8) or (4.5), (4.6) can be written as

$$\xi^1 = \xi^1(u^a, u^{a*}, \underset{0}{\xi^1}, \underset{0}{\xi^{1*}}), \text{ conj.} \tag{4.9}$$

If the u^{a*} actually occur in the right-hand side of (4.9), then these equations define ∞^1 semi-analytic(3) hypersurfaces X_{n-1} in the unitary K_n. We shall show that these u^{a*} do not occur. By forming the total differential of (4.9), we obtain

$$d\xi^1 = du^a\partial_a\xi^1 + du^{a*}\partial_{a*}\xi^1, \text{ conj.} \tag{4.10}$$

Since the du^a, du^{a*} are equal to the differentials $d\xi^\alpha$, $d\xi^{\alpha*}$ $(\alpha=2, \cdots, n)$ of the independent variables ξ^α, $\xi^{\alpha*}$, we find by comparing (4.10) and its conjugate with (4.3), (4.4) that

$$\partial_{a*}\xi^1 = 0, \text{ conj.} \tag{4.11}$$

Hence (4.9) may be written in the form

$$\xi^1 = \xi^1(u^a, \underset{0}{\xi^1}, \underset{0}{\xi^{1*}}), \text{ conj.} \tag{4.12}$$

The equations (4.12) determine ∞^1 analytic hypersurfaces X_{n-1} in the unitary K_n. Such analytic hypersurfaces are always unitary(14) K_{n-1}. Hence our lemma is proved.

In view of the fact that (4.1), (4.2) are unitary orthogonality relations, we state

DEFINITION 2. *The integrals $f=$const. and $f^*=$const. of* (4.1), (4.2) *will be said to define ∞^1 hypersurfaces in the unitary K_n such that the hypersurfaces are completely unitary orthogonal to the congruence vector u_λ.*

We may restate Lemma 2 in terms of Definition 2 as follows:

---

(14) [5, vol. 2, p. 245].

LEMMA 3. *The ∞^1 hypersurfaces which are completely unitary orthogonal to the congruence vector u_λ are ∞^1 unitary K_{n-1} in the unitary K_n.*

Let us now consider the single Pfaffian

$$u_\lambda d\xi^\lambda + u_{\lambda^*} d\xi^{\lambda^*} = 0. \tag{4.13}$$

By assuming that u_1 does not vanish over some domain D of the unitary K_n, we can construct a theory of this Pfaffian in which (4.3) is replaced by

$$d\xi^1 = -\sum{}'(u_\alpha/u_1)d\xi^\alpha - (u_{\lambda^*}/u_1)d\xi^{\lambda^*}, \quad \alpha = 2, \cdots, n; \lambda = 1, \cdots, n. \tag{4.14}$$

The equation (4.5) becomes

$$\xi^1 = \xi^1(\xi^\alpha, \xi^{\lambda^*}, \overset{}{\underset{0}{\xi^1}}). \tag{4.15}$$

Furthermore, (4.13) has only one independent integral

$$\underset{0}{\xi^1} = f(\xi^\lambda, \xi^{\lambda^*}). \tag{4.16}$$

The equation corresponding to (4.4) is identical with (4.14); the equation corresponding to (4.6) is the conjugate of (4.15). Finally, the equation corresponding to (4.8) is equivalent to the conjugate of (4.16). However, this last equation is trivial since if f^* = const. is an integral of (4.13), then $f = F(f^*)$.

We now define a new term.

DEFINITION 3. *The integral f = const. of (4.13) will be said to define ∞^1 hypersurfaces X_{n-1} in the unitary K_n which are semi-unitary orthogonal to the congruence u_λ.*

These semi-unitary orthogonal hypersurfaces can be characterized by their parameter representation. We prove

LEMMA 4. *The ∞^1 semi-unitary orthogonal hypersurfaces to the congruence u_λ cannot possess an analytic parameter representation of rank $(n-1)$. That is, these X_{n-1} are not unitary K_{n-1}.*

Let us assume the contrary, namely, that these hypersurfaces possess an analytic parameter representation.

$$\xi^\lambda = \xi^\lambda(u^a), \qquad \lambda = 1, \cdots, n; a = 1, \cdots, n-1, \text{ conj.} \tag{4.17}$$

Since the rank of (4.17) is $(n-1)$, we can solve for the $(n-1)u^a$ in terms of $(n-1)$ of the ξ^α (say, $\alpha = 2, \cdots, n$),

$$u^a = u^a(\xi^\alpha), \text{ conj.} \tag{4.18}$$

Substituting (4.18) into the first equation of (4.17), we obtain

$$\xi^1 = \xi^1(\xi^\alpha), \text{ conj.} \tag{4.19}$$

Forming the total differential of (4.19), we find

$$(4.20)\qquad d\xi^1 = d\xi^\alpha \partial_\alpha \xi^1, \qquad \alpha = 2, \cdots, n, \text{ conj.}$$

Remembering that ξ^α, ξ^{α^*}, ξ^{1^*} ($\alpha=2, \cdots, n$) are independent variables and comparing with (4.14), we obtain

$$(4.21)\qquad u_{\lambda^*} = 0, \qquad \lambda = 1, \cdots, n, \text{ conj.}$$

Hence the congruence vector vanishes. Thus, the assumption (4.17) is false for a non-vanishing congruence; our lemma is proved. Our lemma implies that if the class of semi-unitary orthogonal hypersurfaces can be parameterized, the parameterization is semi-analytic (see 4.9).

Another relation exists between the ∞^1 completely unitary orthogonal hypersurfaces and the ∞^1 semi-unitary orthogonal hypersurfaces. Let us consider the systems of partial differential equations associated with (4.1), (4.2) and (4.13). The system associated with (4.1), (4.2) is

$$(4.22)\qquad \partial_\alpha f - (u_\alpha/u_1)\partial_1 f = 0, \qquad \alpha = 2, \cdots, n,$$

$$(4.23)\qquad \partial_{\alpha^*} f - (u_{\alpha^*}/u_{1^*})\partial_{1^*} f = 0.$$

The system associated with (4.13) is composed of (4.22), (4.23) plus the additional equation

$$(4.24)\qquad \partial_{1^*} f - (u_{1^*}/u_1)\partial_1 f = 0.$$

If $\partial_1 f$, $\partial_{1^*} f$ do not vanish over the domain D in which u_1, u_{1^*} do not vanish, then the non-vanishing scalars ρ, γ exist such that (4.22), (4.23)—hence (4.1), (4.2)—are equivalent to

$$(4.25)\qquad u_\lambda = \rho \nabla_\lambda f,$$

$$(4.26)\qquad u_{\lambda^*} = \gamma \nabla_{\lambda^*} f.$$

The equation (4.13) is equivalent to (4.25), (4.26) plus the additional equation (4.24). However, the latter implies

$$(4.27)\qquad \gamma = \rho.$$

Before proceeding to enumerate these new results, we note that if $f=$const. is a solution of (4.22) through (4.24), then $f^*=$const. is a solution of the same equations (see the discussion following 4.16). Hence the word "conjugate" can be written after equations (4.25) through (4.27).

Lemma 5. *The solutions $f=$*const. *(and its conjugate) of* (4.25), (4.26) *where $\rho \neq \gamma$ determine ∞^1 completely unitary orthogonal hypersurfaces (unitary K_{n-1}) to the congruence. The solution $f=$*const. *of* (4.25), (4.26) *where $\rho=\gamma$ determine ∞^1 semi-unitary orthogonal hypersurfaces X_{n-1} to the congruence.*

5. Congruences completely unitary orthogonal to $\infty^1 K_{n-1}$ in K_n. We consider congruences which are completely unitary orthogonal to ∞^1 unitary

K_{n-1} in unitary K_n. The integrability conditions(15) of (4.1), (4.2) are

$$u_\lambda \partial_{[\beta} u_{\alpha]} + u_\alpha \partial_{[\lambda} u_{\beta]} + u_\beta \partial_{[\alpha} u_{\lambda]} = 0, \text{ conj.,} \tag{5.1}$$

$$u_\beta \partial_{\alpha^*} u_\lambda - u_\lambda \partial_{\alpha^*} u_\beta = 0, \text{ conj.} \tag{5.2}$$

By use of (2.11), (2.22), we find

$$\partial_{[\beta} u_{\alpha]} = \nabla_{[\beta} u_{\alpha]} + S_{\beta\alpha}^{\cdot\cdot\gamma} u_\gamma, \text{ conj.,} \tag{5.3}$$

$$\partial_{\alpha^*} u_\lambda = \nabla_{\alpha^*} u_\lambda, \text{ conj.} \tag{5.4}$$

Thus (5.1), (5.2) become

$$u_{[\lambda} \nabla_\beta u_{\alpha]} = - u_{[\lambda} S_{\beta\alpha]}^{\cdot\cdot\gamma} u_\gamma, \text{ conj.,} \tag{5.5}$$

$$u_{[\beta} \nabla_{|\alpha^*|} u_{\lambda]} = 0, \text{ conj.} \tag{5.6}$$

By transvecting (5.2) or (5.6) with u^λ, we find

$$\nabla_{\alpha^*} u_\beta = u_\beta (\nabla_{\alpha^*} u_\lambda) u^\lambda, \text{ conj.} \tag{5.7}$$

Replacing the right-hand side by (3.6), we obtain

$$\nabla_{\alpha^*} u_\beta = y_{\alpha^*} u_\beta + q^* u_{\alpha^*} u_\beta, \text{ conj.} \tag{5.8}$$

Hence, upon comparing (5.8) with (3.6), we find that

$$l_{\alpha^*\lambda} = 0, \qquad x_\lambda = 0, \text{ conj.,} \tag{5.9}$$

is a consequence of the integrability conditions (5.6). We now study the meaning of the integrability conditions (5.5). Let us assume that the connection of K_n is semi-symmetric (see 2.23). Then the equation (5.5) reduces to

$$u_{[\lambda} l_{\beta\alpha]} = 0, \text{ conj.,} \tag{5.10}$$

in consequence of (3.5) and (5.5). By transvecting (5.10) with u^λ and using (3.2), (3.3), this equation becomes

$$l_{[\beta\alpha]} = 0, \text{ conj.} \tag{5.11}$$

Conversely, if (5.9) and (5.11) are valid and if the connection of K_n is semi-symmetric, then the expressions (3.5), (3.6) satisfy the integrability conditions (5.5), (5.6). This leads us to

THEOREM 2. *Consider a unitary space K_n with semi-symmetric connection and such that the solutions u_λ of* (3.5), (3.6) *exist then if and only if*: (1) *$l_{\alpha\beta}$ is symmetric*; (2) *$l_{\alpha^*\beta}$, x_λ vanish, does the vector u_λ define a congruence which is completely unitary orthogonal to ∞^1 hypersurfaces in the unitary K_n.*

(15) [4, p. 29, equation 23]. Since the u_λ, u_{λ^*} in (4.1), (4.2) are functions of ξ^μ and ξ^{μ^*}, the *complete* Pfaffians in (4.1), (4.2) can be written as $w_{\lambda^*} d\xi^{\lambda^*} + u_\lambda d\xi^\lambda = 0$, $u_{\lambda^*} d\xi^{\lambda^*} + w_\lambda d\xi^\lambda = 0$, where $w_\lambda, w_{\lambda^*} = 0$. If one writes out the equation 23, p. 29 of [4], then for unstarred variables (or indices) the equation (5.1) results; if one of the starred variables (or indices) is used then equation (5.2) results.

Let us now restrict ourselves to real congruences (see Definition 1 (a)). By introducing the Frenet formulas[18] for the ∞^{n-1} curves X_1 of the congruence, we can determine the meaning of the vector w_α in (3.5). From the Frenet formulas, it follows that

$$u^{\alpha}\nabla_{\alpha}u_{\lambda} + u^{\alpha^*}\nabla_{\alpha^*}u_{\lambda} = \underset{00}{k}\, u_{\lambda} + \underset{01}{k}\, \underset{1}{u}_{\lambda}, \text{ conj.}, \tag{5.12}$$

where k (subindex 00, 01) are curvatures and u_λ (subindex 1) is the first normal of each X_1 in the unitary K_n. By use of (3.5), (3.6), (3.13), we find

$$u^{\alpha}\nabla_{\alpha}u_{\lambda} + u^{\alpha^*}\nabla_{\alpha^*}u_{\lambda} = w_{\lambda} + x_{\lambda} + (p - p^*)u_{\lambda}, \text{ conj.} \tag{5.13}$$

By comparison of (5.12), (5.13), we obtain

$$w_{\lambda} + x_{\lambda} = \underset{01}{k}\, \underset{1}{u}_{\lambda}, \text{ conj.}, \tag{5.14}$$

$$p - p^* = \underset{00}{k}, \text{ conj.} \tag{5.15}$$

If we require that the congruence shall be completely unitary orthogonal to ∞^1 hypersurfaces in the unitary K_n, then it follows from Theorem 2 that the vector x_λ vanishes. Hence, we obtain the result

THEOREM 3. *If the congruence is real and completely unitary orthogonal to ∞^1 hypersurfaces in the unitary K_n, then* (1) *the vector w_λ lies along the first normal to any X_1 of the congruence*; (2) *the magnitude of w_λ is equal to the* (0, 1) *curvature of X_1*; (3) *the imaginary part of the scalar p is one-half the* (0, 0) *curvature of X_1.*

Again, let us consider the case where the congruence is completely unitary orthogonal to ∞^1 hypersurfaces in the unitary K_n and where the congruence is either real or complex Euclidean. By means of Theorems 1 and 2, we may write (3.5), (3.6) in the form

$$\nabla_{\alpha}u_{\lambda} = l_{\alpha\lambda} + u_{\alpha}w_{\lambda} + h^{\cdot}_{\alpha}u_{\lambda}, \text{ conj.}, \tag{5.16}$$

$$\nabla_{\alpha^*}u_{\lambda} = -\, h_{\alpha^*}u_{\lambda}, \text{ conj.}, \tag{5.17}$$

where

$$h_{\alpha} = z_{\alpha} + pu_{\alpha}, \text{ conj.} \tag{5.18}$$

From (5.17), we see that if $\nabla_{\alpha^*}u_\lambda$ vanishes, then h_α vanishes. Hence from (5.18), we find that the vector z_α and the scalar p vanish. By (5.15), this last result means that the curvature k (subindex 00) of a real congruence vanishes. Furthermore, the equations (5.16) and (5.11) furnish the result

$$\nabla_{\alpha}u_{\lambda} = l_{\alpha\lambda} + u_{\alpha}w_{\lambda}, \text{ conj.} \tag{5.19}$$

---

(18) [2, equation 3.23].

The equation (5.19) is the condition([8]) satisfied by a congruence of curves V_1 which are orthogonal to ∞^1 hypersurfaces in a Riemannian space of n dimensions V_n. Hence, we have the result

THEOREM 4. *If the vector u_λ determines a congruence which is*: (1) *completely unitary orthogonal to ∞^1 hypersurfaces in the unitary K_n*; (2) *the vector u_λ is analytic*; (3) *the congruence is either real or complex Euclidean, then the conditions satisfied by the congruence vector u_λ in the unitary K_n are identical with those satisfied by the congruence vector u_λ which is orthogonal to ∞^1 hypersurfaces in V_n. If the condition* (3) *is replaced by the stronger requirement that the congruence is real, then a further conclusion follows. Namely, the curvature k (subindex* 00) *vanishes.*

6. **Congruences semi-unitary orthogonal to $\infty^1 X_{n-1}$ in K_n.** If the congruence is semi-unitary orthogonal to ∞^1 hypersurfaces in the unitary K_n, then the integrability conditions of (4.13) must be satisfied. These are given by

$$(6.1)\qquad u_\lambda \partial_{[\beta} u_{\alpha]} + u_\alpha \partial_{[\lambda} u_{\beta]} + u_\beta \partial_{[\alpha} u_{\lambda]} = 0, \text{ conj.},$$

$$(6.2)\qquad u_\lambda \partial_{[\beta^*} u_{\alpha]} + u_\alpha \partial_{[\lambda} u_{\beta^*]} + u_{\beta^*} \partial_{[\alpha} u_{\lambda]} = 0, \text{ conj.}$$

By use of (5.3), (5.4), the equations (6.1), (6.2) become

$$(6.3)\qquad u_{[\lambda} \nabla_\beta u_{\alpha]} = - u_{[\lambda} S_{\beta\alpha]}^{\cdot\cdot\gamma} u_\gamma, \text{ conj.},$$

$$(6.4)\qquad u_{[\lambda} \nabla_{\beta^*} u_{\alpha]} = - (1/3) u_{\beta^*} S_{\alpha\lambda}^{\cdot\cdot\gamma} u_\gamma, \text{ conj.}$$

Let us assume that the connection of K_n is semi-symmetric (see 2.23). Then, the right-hand side of (6.3) vanishes and the right-hand side of (6.4) becomes $(-(1/3)u_{\beta^*}u_{[\alpha}p_{\lambda]})$. Upon substituting (3.5) into (6.3), we obtain

$$(6.5)\qquad u_{[\lambda} l_{\beta\alpha]} = 0, \text{ conj.}$$

Transvecting with u^λ and using (3.2), (3.3), we conclude that

$$(6.6)\qquad l_{[\beta\alpha]} = 0, \text{ conj.}$$

Conversely, if (6.6) is valid, then (3.5) satisfies (6.3). We next study the consequences of (6.4). By transvecting (6.4) with u^α, u^{β^*}, we obtain

$$(6.7)\qquad u_\lambda u^\alpha \nabla_{[\beta^*} u_{\alpha]} + \nabla_{[\lambda} u_{\beta^*]} + u_{\beta^*} u^\alpha \nabla_{[\alpha} u_{\lambda]} = - u_{\beta^*} u^\alpha u_{[\alpha} p_{\lambda]}, \text{ conj.},$$

$$(6.8)\qquad u_\lambda u^{\beta^*} \nabla_{[\beta^*} u_{\alpha]} + u_\alpha u^{\beta^*} \nabla_{[\lambda} u_{\beta^*]} + \nabla_{[\alpha} u_{\lambda]} = - u_{[\alpha} p_{\lambda]}, \text{ conj.}$$

Due to the symmetry of (6.4), no additional relations are obtained by further transvection with u^λ. With the aid of (3.5), (3.6), (6.6), the two previous equations become

$$(6.9)\qquad \begin{aligned} u_\lambda u^\alpha (u_{[\beta^*} x_{\alpha]} + y_{[\beta^*} u_{\alpha]}) &+ l_{[\lambda\beta^*]} + u_{[\lambda} x_{\beta^*]} + y_{[\lambda} u_{\beta^*]} \\ &+ u_{\beta^*} u^\alpha (u_{[\alpha} w_{\lambda]} + z_{[\alpha} u_{\lambda]}) = - u_\beta u^\alpha u_{[\alpha} p_{\lambda]}, \text{ conj.}, \end{aligned}$$

$$(6.10)\qquad \begin{aligned} u_\lambda u^{\beta^*} (u_{[\beta^*} x_{\alpha]} + y_{[\beta^*} u_{\alpha]}) &+ u_\alpha u^{\beta^*} (u_{[\lambda} x_{\beta^*]} + y_{[\lambda} u_{\beta^*]}) \\ &+ (u_{[\alpha} w_{\lambda]} + z_{[\alpha} u_{\lambda]}) = - u_{[\alpha} p_{\lambda]}, \text{ conj.} \end{aligned}$$

Simplifying with the aid of (3.4), we find

$$(6.11)\quad 2l_{[\lambda\beta^{\cdot}]} - u_{\beta^{\cdot}}(x_\lambda + z_\lambda - y_\lambda - w_\lambda) = - u_{\beta^{\cdot}}[p_\lambda - u_\lambda(p_\alpha u^\alpha)], \text{ conj.},$$

$$(6.12)\quad - u_\alpha(x_\lambda + z_\lambda - y_\lambda - w_\lambda - p_\lambda) + u_\lambda(x_\alpha + z_\alpha - y_\alpha - w_\alpha - p_\alpha) = 0, \text{conj.}$$

Transvecting (6.11) with $u^{\beta^{\cdot}}$ and simplifying with (3.2), (3.3), we obtain

$$(6.13)\quad x_\lambda + z_\lambda - y_\lambda - w_\lambda = p_\lambda - u_\lambda(p_\alpha u^\alpha), \text{ conj.}$$

Substituting (6.13) into (6.11), we find

$$(6.14)\quad l_{[\lambda\beta^{\cdot}]} = 0, \text{ conj.}$$

By substituting (6.13) into (6.12), we find that the latter equation is identically satisfied. Thus (6.13), (6.14) are the only equations obtained by transvecting (6.4) with u^λ. Conversely, by expanding (6.4) and using (3.5), (3.6), we find that in virtue of (6.6), (6.14), (6.13), the equation (6.4) is identically satisfied. This leads us to a theorem which is similar to Theorem 2,

THEOREM 5. *Consider a unitary space K_n with semi-symmetric connection and such that the solutions u_λ of* (3.5), (3.6) *exist, then if and only if*: (1) $l_{\alpha\beta}$, $l_{\alpha^{\cdot}\beta}$ *are symmetric*; (2) $x_\lambda + z_\lambda - y_\lambda - w_\lambda = p_\lambda - u_\lambda(p_\alpha u^\alpha)$, *does the vector u_λ define a congruence which is semi-unitary orthogonal to ∞^1 hypersurfaces in the unitary K_n.*

By use of Lemma 5, we can obtain some further properties of the semi-unitary congruences. From (4.25), (4.26), (4.27) it follows that for congruences which are semi-unitary orthogonal to ∞^1 hypersurfaces in the unitary K_n

$$(6.15)\quad u_\lambda = \rho\nabla_\lambda f, \text{ conj.},$$

$$(6.16)\quad u_{\lambda^{\cdot}} = \rho\nabla_{\lambda^{\cdot}} f, \text{ conj.}$$

By covariant differentiation of (6.15), (6.16), we find

$$(6.17)\quad \nabla_\alpha u_\lambda = \rho\nabla_\alpha\nabla_\lambda f + (\nabla_\lambda f)(\nabla_\alpha \rho), \text{ conj.},$$

$$(6.18)\quad \nabla_\alpha u_{\lambda^{\cdot}} = \rho\nabla_\alpha\nabla_{\lambda^{\cdot}} f + (\nabla_{\lambda^{\cdot}} f)(\nabla_\alpha \rho), \text{ conj.}$$

From the relations (2.11), (2.22), we see that

$$(6.19)\quad \nabla_\alpha\nabla_\lambda f = \nabla_\lambda\nabla_\alpha f + 2S_{\lambda\alpha}^{\cdot\cdot\gamma}\nabla_\gamma f, \text{ conj.},$$

$$(6.20)\quad \nabla_\alpha\nabla_{\lambda^{\cdot}} f = \nabla_{\lambda^{\cdot}}\nabla_\alpha f, \text{ conj.}$$

Substituting the last two equations into (6.17), (6.18), the latter become

$$(6.21)\quad \nabla_\alpha u_\lambda = \rho\nabla_\lambda\nabla_\alpha f + 2\rho S_{\lambda\alpha}^{\cdot\cdot\gamma}\nabla_\gamma f + (\nabla_\lambda f)(\nabla_\alpha \rho), \text{ conj.},$$

$$(6.22)\quad \nabla_\alpha u_{\lambda^{\cdot}} = \rho\nabla_{\lambda^{\cdot}}\nabla_\alpha f + (\nabla_{\lambda^{\cdot}} f)(\nabla_\alpha \rho) \text{ conj.},$$

Simplifying (6.21), (6.22) by use of the equations (6.15) through (6.18), we

obtain

$$\nabla_\alpha u_\lambda = \nabla_\lambda u_\alpha + 2S_{\lambda\alpha}^{\cdot\cdot\gamma} u_\gamma + 2\rho^{-1} u_{[\lambda} \nabla_{\alpha]}\rho, \text{ conj.,} \tag{6.23}$$

$$\nabla_\alpha u_{\lambda^*} = \nabla_{\lambda^*} u_\alpha + 2\rho^{-1} u_{[\lambda^*} \Delta_{\alpha]}\rho, \text{ conj.} \tag{6.24}$$

By use of (3.5), (3.6), (6.6), (6.14), the above two equations become

$$u_{[\alpha} w_{\lambda]} + z_{[\alpha} u_{\lambda]} = S_{\lambda\alpha}^{\cdot\cdot\gamma} u_\gamma + \rho^{-1} u_{[\lambda} \nabla_{\alpha]}\rho, \text{ conj.,} \tag{6.25}$$

$$u_{[\alpha} x_{\lambda^*]} + y_{[\alpha} u_{\lambda^*]} = \rho^{-1} u_{[\lambda^*} \nabla_{\alpha]}\rho, \text{ conj.} \tag{6.26}$$

Let us assume that the connection of K_n is semi-symmetric. Transvecting the previous two equations with u^α, we obtain

$$w_\lambda - z_\lambda = - p_\lambda + u_\lambda(p_\alpha u^\alpha) - \nabla_\lambda \ln \rho + u_\lambda(u^\alpha \nabla_\alpha \ln \rho), \text{ conj.,} \tag{6.27}$$

$$x_{\lambda^*} - y_{\lambda^*} = - \nabla_{\lambda^*} \ln \rho + u_{\lambda^*}(u^\alpha \nabla_\alpha \ln \rho), \text{ conj.} \tag{6.28}$$

We are now in a position to prove

THEOREM 6. *If*: (1) *the connection of the unitary K_n is semi-symmetric*; (2) *the congruence is semi-unitary orthogonal to ∞^1 hypersurfaces in K_n*; (3) *the congruence is real or complex Euclidean*; (4) $w_\lambda = z_\lambda$, $x_\lambda = y_\lambda$, *then every two hypersurfaces X_{n-1} intercept equal arc segments on all curves of the congruence.*

From condition (4) of our theorem, that is,

$$w_\lambda = z_\lambda, \qquad x_\lambda = y_\lambda, \text{ conj.,} \tag{6.29}$$

it follows by use of (6.13) that

$$p_\lambda = (p_\alpha u^\alpha) u_\lambda, \text{ conj.} \tag{6.30}$$

Substituting (6.29), (6.30) into (6.27), (6.28), we find

$$u_\lambda = \theta \nabla_\lambda \rho, \text{ conj.,} \tag{6.31}$$

$$u_{\lambda^*} = \theta \nabla_{\lambda^*} \rho, \text{ conj.,} \tag{6.32}$$

where θ is some function of ξ^λ, ξ^{λ^*}. Thus ρ is an integral of the system (4.22) through (4.24). Since that system has only one independent integral, namely, $f(\xi^\lambda, \xi^{\lambda^*})$, it follows that

$$\rho = F(f), \text{ conj.,} \tag{6.33}$$

where $F(f)$ is some arbitrary function of f. From (6.33), (6.15), (6.16), we find

$$F(f) df = u_\lambda d\xi^\lambda + u_{\lambda^*} d\xi^{\lambda^*}, \tag{6.34}$$

for arbitrary $d\xi^\lambda$, $d\xi^{\lambda^*}$. Now let us consider the vector $(d\xi^\lambda, d\xi^{\lambda^*})$ as in the direction of u^λ. By multiplying and dividing the right-hand side of (6.34) by ds (the element of arc length along a curve of the congruence), we obtain

$$F(f) df = 2ds. \tag{6.35}$$

Integrating (6.35) between $f=c_0, f=c_1$, we find

$$2(s-s_0)=\int_{c_0}^{c_1}F(f)df. \tag{6.36}$$

The fact that the right-hand side of (6.36) is independent of any particular curve of the congruence proves our theorem.

We can obtain the geometric meaning of the essential condition (4) in Theorem 6 by limiting ourselves to real congruences. We prove

LEMMA 6. *Consider a real congruence which is semi-unitary orthogonal to* ∞^1 *hypersurfaces in a unitary* K_n *with semi-symmetric connection, then if and only if*: (1) *the* (01) *curvature of each* X_1 *vanishes*; (2) $p_\lambda=(p_\alpha u^\alpha)u_\lambda$, *are the equations* (6.29) *valid.*

First, we show the sufficiency of our conditions. From the first condition, we have

$$\underset{01}{k}=0,\ \text{conj.} \tag{6.37}$$

Hence from (5.14), it follows that

$$w_\lambda+x_\lambda=0,\ \text{conj.} \tag{6.38}$$

Since the congruence is real, the equation (3.12), is valid, that is,

$$z_\lambda=-y_\lambda,\ \text{conj.} \tag{6.39}$$

By use of the second condition and (6.13), we obtain

$$x_\lambda+z_\lambda-y_\lambda-w_\lambda=0,\ \text{conj.} \tag{6.40}$$

By substituting (6.38), (6.39) into (6.40), we obtain the equations (6.29).

Conversely, from (6.29) and the fact that the congruence is real, it follows that the conditions (1), (2) of our theorem are satisfied. Thus, from (6.29) and (6.13), we obtain

$$p_\lambda=(p_\alpha u^\alpha)u_\lambda,\ \text{conj.} \tag{6.41}$$

Furthermore, by use of (6.39), (6.29) and (5.14), we find that k (subindex 01) vanishes.

We now translate Theorem 6 into terms connected with a real congruence of geodesic curves. Our result is

THEOREM 7. *If the curves* X_1 *of a real congruence in a unitary space* K_n *with semi-symmetric connection satisfy the conditions*: (1) *the congruence is semi-unitary orthogonal to* ∞^1 *hypersurfaces in* K_n; (2) *the* ∞^{n-1} X_1 *are geodesic*; *and either* (3) *the curvature* k (*subindex* 01) *vanishes*; *or* (4) $p_\lambda=(p_\alpha u^\alpha)u_\lambda$, *then every two hypersurfaces intercept equal arc segments on all* X_1 *of the congruences.*

It has been shown(17) that an X_1 in a unitary K_n with semi-symmetry connection is geodesic if and only if

$$p_\lambda = (p_\alpha u^\alpha)u_\lambda - \underset{01}{k}\,\underset{1}{u_\lambda}, \text{ conj.,} \tag{6.42}$$

$$\underset{00}{k} = 0, \text{ conj.} \tag{6.43}$$

The condition (6.43) is of no use to us. From (6.42), we note that if condition (3) of our theorem is valid, then the condition (4) is necessarily satisfied, and conversely. From Lemma 6, it follows that (6.29) is valid. Hence, Lemma 6 leads to the desired conclusion.

If in particular, the space K_n has a symmetric connection, then

$$p_\lambda = 0, \text{ conj.} \tag{6.44}$$

Thus the conditions (3), (4) of Theorem 7 are satisfied. Theorem 7 becomes analogous to a theorem in Riemannian space(4).

References

1. N. Coburn, *Unitary curves in unitary space*, Revista de la Universidad Nacional de Tucuman vol. 2 (1941) pp. 159–167.

2. ———, *Frenet formulas for curves in unitary space*, Journal of Mathematics and Physics, Massachusetts Institute of Technology vol. 21 (1942) pp. 10–18.

3. ———, *Semi-analytic unitary subspaces of unitary space*, Amer. J. Math. vol. 64 (1942) pp. 714–724.

4. T. Levi-Civita, *The absolute differential calculus*, London, 1929.

5. J. A. Schouten and D. J. Struik, *Einfuhrung in die Neueren Methoden der Differentialgeometrie*, Groningen, vol. 1, 1935; vol. 2, 1938.

University of Texas,
Austin, Texas

(17) [2, Theorem 4].

RECURSIVE PREDICATES AND QUANTIFIERS(1)

BY
S. C. KLEENE

This paper contains a general theorem on the quantification of recursive predicates, with applications to the foundations of mathematics. The theorem (Theorem II) is a slight extension of previous results on Herbrand-Gödel general recursive functions(2), while the applications include theorems of Church (Theorem VII)(3) and Gödel (Theorem VIII)(4) and other incompleteness theorems. It is thought that in this treatment the relationship of the results stands out more clearly than before.

The general theorem asserts that to each of an enumeration of predicate forms, there is a predicate not expressible in that form. The predicates considered belong to elementary number theory.

The possibility that this theorem may apply appears whenever it is proposed to find a necessary and sufficient condition of a certain kind for some given property of natural numbers; in other words, to find a predicate of a given kind equivalent to a given predicate. If the specifications on the predicate which is being sought amount to its having one of the forms listed in the theorem, then for some selection of the given property a necessary and sufficient condition of the desired kind cannot exist.

In particular, it is recognized that to find a complete algorithmic theory for a predicate $P(a)$ amounts to expressing the predicate as a recursive predicate. By one of the cases of the theorem, this is impossible for a certain $P(a)$, which gives us Church's theorem.

Again, when we recognize that to give a complete formal deductive theory (symbolic logic) for a predicate $P(a)$ amounts to finding an equivalent predicate of the form $(Ex)R(a, x)$ where $R(a, x)$ is recursive, we have immediately Gödel's theorem, as another case of the general theorem.

Still another application is made, when we consider the nature of a constructive existence proof. It appears that there is a proposition provable classically for which no constructive proof is possible (Theorem X).

The endeavor has been made to include a fairly complete exposition of definitions and results, including relevant portions of previous theory, so that

Presented to the Society, September 11, 1940; received by the editors February 13, 1942. In the abstract of this paper, Bull. Amer. Math. Soc. abstract 46-11-464, erratum: line 4, for "for." read "for all.".

(1) A part of the work reported in this paper was supported by the Institute for Advanced Study and the Alumni Research Foundation of the University of Wisconsin.

(2) Gödel [2, §9] (see the bibliography at the end of the paper).

(3) Church [1].

(4) Gödel [1, Theorem VI].

the paper should be self-contained, although some details of proof are omitted.

The general theorem is obtained quickly in Part I from the properties of the μ-operator, or what essentially was called the $\mathfrak{p}$-function in the author's dissertation([8]). Part II contains some variations on the theme of Part I, and may be omitted by the cursory reader. The applications to foundational questions are in Part III, only a few passages of which depend on Part II.

I. The general theorem on recursive predicates and quantifiers

1. Primitive recursive functions. The discussion belongs to the context of the informal theory of the natural numbers

$$0, 1, 2, \cdots, x, x', \cdots.$$

The functions which concern us are number-theoretic functions, for which the arguments and values are natural numbers.

We consider the following schemata as operations for the definition of a function ϕ from given functions appearing in the right members of the equations (c is any constant natural number):

$$\text{(I)} \qquad \phi(x) = x',$$

$$\text{(II)} \qquad \phi(x_1, \cdots, x_n) = c,$$

$$\text{(III)} \qquad \phi(x_1, \cdots, x_n) = x_i,$$

$$\text{(IV)} \qquad \phi(x_1, \cdots, x_n) = \theta(\chi_1(x_1, \cdots, x_n), \cdots, \chi_m(x_1, \cdots, x_n)),$$

$$\text{(Va)} \qquad \begin{cases} \phi(0) = c \\ \phi(y') = \chi(y, \phi(y)), \end{cases}$$

$$\text{(Vb)} \qquad \begin{cases} \phi(0, x_1, \cdots, x_n) = \psi(x_1, \cdots, x_n) \\ \phi(y', x_1, \cdots, x_n) = \chi(y, \phi(y, x_1, \cdots, x_n), x_1, \cdots, x_n). \end{cases}$$

Schema (I) introduces the successor function, Schema (II) the constant functions, and Schema (III) the identity functions. Schema (IV) is the schema of definition by substitution, and Schema (V) the schema of primitive recursion. Together we may call them (and more generally, schemata reducible to a series of applications of them) the *primitive recursive* schemata.

A function ϕ which can be defined from given functions $\psi_1, \cdots, \psi_k$ by a series of applications of these schemata we call *primitive recursive* in the given functions; and in particular, a function ϕ definable ab initio by these means, *primitive recursive*.

Now let us consider number-theoretic predicates, that is, propositional functions of natural numbers.

([8]) Kleene [1, §18].

In asserting propositions, and in designating predicates, we use a logical symbolism, as follows. Operations of the propositional calculus: & (and), $\vee$ (or), $^{-}$ (not), $\rightarrow$ (implies), $\equiv$ (equivalent). Quantifiers: (x) (for all x), (Ex) (there exists an x such that). These operations may be taken either in the sense of classical mathematics, or in the sense of constructive or intuitionistic mathematics, except where one or the other of the two interpretations is specified.

A predicate $P(x_1, \cdots, x_n)$ is said to be ***primitive recursive***, if there is a primitive recursive function $\pi(x_1, \cdots, x_n)$ such that

$$(1) \qquad P(x_1, \cdots, x_n) \equiv \pi(x_1, \cdots, x_n) = 0.$$

We can without loss of generality restrict π to take only 0 and 1 as values, and call it in this case the ***representing function*** of P.

Under classical interpretations, which give a dichotomy of propositions into true and false, we can assign to any predicate P a ***representing function*** π which has 0 or 1 as value according as the value of P is true or false; and then say that P is ***primitive recursive*** if π is.

2. **General recursive functions.** We shall proceed to the Herbrand-Gödel generalization of the notion of recursive function. We start with a preliminary account, certain features of which we shall then restate carefully.

The way in which the function ϕ is defined from the given functions in an application of one of the primitive recursive schemata amounts to this: the values $\phi(x_1, \cdots, x_n)$ of ϕ for the various sets $x_1, \cdots, x_n$ of arguments are determined unambiguously by the equations and the values of the given functions, using only principles of determination which we can formalize as a substitution rule and a replacement rule.

The formalization presupposes suitable conventions governing the symbolism, which are easily supplied. In particular, we must distinguish between the ***variables*** for numbers and the ***numerals***, that is the expressions for the fixed numbers in terms of the symbols for 0 and the successor operation $'$. The rules are the following.

R1: ***to substitute, for the variables*** $\mathrm{x}_1, \cdots, \mathrm{x}_n$ ***of an equation, numerals*** $\mathbf{x}_1, \cdots, \mathbf{x}_n$***, respectively.***

R2: ***to replace a part*** $\mathrm{f}(\mathbf{x}_1, \cdots, \mathbf{x}_n)$ ***of the right member of an equation by*** $\mathbf{x}$***, where*** f ***is a function symbol, where*** $\mathbf{x}_1, \cdots, \mathbf{x}_n, \mathbf{x}$ ***are numerals, and where*** $\mathrm{f}(\mathbf{x}_1, \cdots, \mathbf{x}_n) = \mathbf{x}$ ***is a given equation.***

By a ***given*** equation $\mathrm{f}(\mathbf{x}_1, \cdots, \mathbf{x}_n) = \mathbf{x}$ for R2, we mean an equation expressing one of the values of one of the given functions for the schema application, or an equation of this form already derived by R1 and R2 from the equations of the schema application.

Now let us consider any operation or schema, for the definition of a function in terms of given functions, which can be expressed by a system of equations determining the function values in this manner. In general the equations shall be allowed to contain, besides the *principal* function symbol which represents the function defined, and the *given* function symbols which represent the given functions, also *auxiliary* function symbols. The given function symbols shall not appear in the left members of the equations. Such a schema we shall call *general recursive.*

A function ϕ which can be defined from given functions $\psi_1, \cdots, \psi_k$ by a series of applications of general recursive schemata we call *general recursive* in the given functions; and in particular, a function ϕ definable ab initio by these means we call *general recursive.*

Suppose that a function ϕ is defined, either from given functions $\psi_1, \cdots, \psi_k$ or ab initio, by a succession of general recursive operations. Let us combine the successive systems of equations which effect the definition into one system, using different symbols as principal and auxiliary function symbols in each of the successive systems, and in the resulting system considering as auxiliary all of the function symbols but that representing ϕ and those representing $\psi_1, \cdots, \psi_k$. The restriction imposed on a general recursive schema that the given function symbols should not appear on the left will prevent any ambiguity being introduced by the interaction under R1 and R2 of equations in the combined system which were formerly in separate systems. Thus the definition can be considered as effected in a single general recursive operation.

In particular, any general recursive function can be defined ab initio in one operation, so that in the defining equations there are no given function symbols and what we have called the given equations for an application of R2 must all be derivable from the defining equations by previous applications of R1 and R2. For the formal development, it is convenient to adopt the convention that the principal function symbol shall be that one of the function symbols occurring in the equations of the system which comes latest in a preassigned list of function symbols. The function is then completely described by giving the system of defining equations.

We now restate the definition of general recursive function from this point of view.

A function $\phi(x_1, \cdots, x_n)$ is *GENERAL RECURSIVE*, if there is a system E of equations which defines it recursively in the following sense. A system E of equations *defines recursively* a GENERAL RECURSIVE function of n variables if, for each set $x_1, \cdots, x_n$ of natural numbers, an equation of the form $\mathrm{f}(\mathbf{x}_1, \cdots, \mathbf{x}_n) = \mathbf{x}$, where f is the principal function symbol of E, and where $\mathbf{x}_1, \cdots, \mathbf{x}_n$ are the numerals representing the natural numbers $x_1, \cdots, x_n$, is derivable from E by R1 and R2 for EXACTLY one numeral $\mathbf{x}$. The function of n variables which is defined by E in this case is the func-

tion ϕ, of which the value $\phi(x_1, \cdots, x_n)$ for $x_1, \cdots, x_n$ as arguments is THE NATURAL NUMBER x REPRESENTED BY THE NUMERAL x.

A predicate $P(x_1, \cdots, x_n)$ is ***general recursive***, if there is a general recursive function $\pi(x_1, \cdots, x_n)$ taking only 0 and 1 as values such that (1) holds; in this case, π is called the ***representing function*** of P. (Or, if we introduce the ***representing function*** π first, P is ***general recursive*** if π is.)

3. **The μ-operator.** Consider the operator: μy (the least y such that). If this operator is applied to a predicate $R(x_1, \cdots, x_n, y)$ of the $n+1$ variables $x_1, \cdots, x_n, y$, and if this predicate satisfies the condition

$$(x_1) \cdots (x_n)(Ey)R(x_1, \cdots, x_n, y), \tag{2}$$

we obtain a function $\mu y R(x_1, \cdots, x_n, y)$ of the remaining n free variables $x_1, \cdots, x_n$.

Thence we have a new schema,

$$\phi(x_1, \cdots, x_n) = \mu y[\rho(x_1, \cdots, x_n, y) = 0], \tag{VI$_1$}$$

for the definition of a function ϕ from a given function ρ which satisfies the condition

$$(x_1) \cdots (x_n)(Ey)[\rho(x_1, \cdots, x_n, y) = 0]. \tag{3}$$

We now show that this schema, subject to the condition on ρ, is, like (I)–(V), general recursive. For this purpose, we rewrite it in terms of equations, using an auxiliary function symbol "σ":

$$\text{(VI}_2\text{)} \quad \begin{cases} \sigma(0, x_1, \cdots, x_n, y) = y \\ \sigma(z', x_1, \cdots, x_n, y) = \sigma(\rho(x_1, \cdots, x_n, y'), x_1, \cdots, x_n, y') \\ \qquad \phi(x_1, \cdots, x_n) = \sigma(\rho(x_1, \cdots, x_n, 0), x_1, \cdots, x_n, 0). \end{cases}$$

Assuming the values of ρ, these equations will lead us to the values of ϕ as defined by (VI$_1$), and to only those values, as follows.

Consider informally any fixed set of values of $x_1, \cdots, x_n$ (formally, this means to substitute the corresponding set of numerals for the variables "x_1", $\cdots$, "x_n"). We seek to obtain the corresponding value of $\phi(x_1, \cdots, x_n)$ by replacements on the third equation, and this is the only possibility we have for obtaining that value under the two principles. First we can replace $\rho(x_1, \cdots, x_n, 0)$ by its value, and this is the only first replacement step possible on that equation. According as that value is 0 or is not 0, we seek the value of σ for the next replacement step from the first or second of the equations, and this is the only possible source for the next replacement value. In the first case, we obtain 0 as that value; in the second, we use the value of $\rho(x_1, \cdots, x_n, 1)$ in the second equation, and then seek another value of σ. We continue thus, with no choice in the procedure at any stage. The first case

is first encountered when we come to use the value of $\rho(x_1, \cdots, x_n, y)$ for the first y for which that value is 0, and hence certainly for at most the y given by (3). When this happens, we can complete the pending replacements to obtain that y as the value of $\phi(x_1, \cdots, x_n)$. Thus we get the intended value; and because we had no choice at any stage of the procedure, we can get no other value.

The general recursiveness of the new schema is thus established. Hence, if $R(x_1, \cdots, x_n, y)$ is a general recursive predicate and (2) holds, by taking as ρ the representing function of R, we can conclude that $\mu y R(x_1, \cdots, x_n, y)$ is a general recursive function.

What can we conclude if (2) is not assumed to hold? In this case, $\mu y R(x_1, \cdots, x_n, y)$ may not be completely defined as a function of the variables $x_1, \cdots, x_n$; but for any fixed set of values of $x_1, \cdots, x_n$, the sequence of steps by which we attempt to determine a value for $\phi(x_1, \cdots, x_n)$ from the equations remains as described for the preceding case, only with now the matter of its termination in doubt. If $(Ey)R(x_1, \cdots, x_n, y)$ does hold for that set of values of $x_1, \cdots, x_n$, then it does terminate as described, with $\mu y R(x_1, \cdots, x_n, y)$ as the value; while conversely, if it does terminate, this can only be in consequence of a 0 being encountered among the values of $\rho(x_1, \cdots, x_n, y)$, so that $(Ey)R(x_1, \cdots, x_n, y)$ does hold, and $\mu y R(x_1, \cdots, x_n, y)$ is the value.

Hence, in formal terms, if F is the system of equations obtained by adjoining, to any system E which defines ρ recursively, equations of the form (VI_2), with the notation so arranged that "ϕ" becomes the principal function symbol f, then: an equation of the form $\mathrm{f}(\mathbf{x}_1, \cdots, \mathbf{x}_n) = \mathbf{x}$, where $\mathbf{x}_1, \cdots, \mathbf{x}_n$ are the numerals representing the natural numbers $x_1, \cdots, x_n$, and where $\mathbf{x}$ is a numeral, is derivable from F by R1 and R2 if and only if $(Ey)R(x_1, \cdots, x_n, y)$.

4. **The enumeration theorem.** We introduce a metamathematical predicate $\mathfrak{S}_n$ (for each particular n) as follows.

$\mathfrak{S}_n(\mathrm{Z}, x_1, \cdots, x_n, \mathrm{Y})$: Z *is a system of equations, and* Y *is a formal deduction from* Z *by* R1 *and* R2 *of an equation of the form* $\mathrm{f}(\mathbf{x}_1, \cdots, \mathbf{x}_n) = \mathbf{x}$, *where* f *is the principal function symbol of* Z, *where* $\mathbf{x}_1, \cdots, \mathbf{x}_n$ *are the numerals representing the natural numbers* $x_1, \cdots, x_n$, *and where* $\mathbf{x}$ *is a numeral.*

With this notation, we can state the last result of the preceding section symbolically:

$$(4) \qquad (Ey)R(x_1, \cdots, x_n, y) \equiv (E\mathrm{Y})\mathfrak{S}_n(\mathrm{F}, x_1, \cdots, x_n, \mathrm{Y}).$$

From a like exploration of the possibility that the sequence of steps does not terminate, or simply from (4) by contraposition, we have also:

$$(5) \qquad (y)\overline{R}(x_1, \cdots, x_n, y) \equiv (\mathrm{Y})\overline{\mathfrak{S}}_n(\mathrm{F}, x_1, \cdots, x_n, \mathrm{Y}).$$

Using Gödel's idea of arithmetizing metamathematics(6), suppose that natural numbers have been correlated to the formal objects, distinct numbers to distinct objects. The metamathematical predicate $\mathfrak{S}_n(Z, x_1, \cdots, x_n, Y)$ is carried by the correlation into a number-theoretic predicate $S_n(z, x_1, \cdots, x_n, y)$, the definition of which we complete by taking it as false for values of z, y not both correlated to formal objects.

For a suitably chosen Gödel numbering, we can show, with a little trouble that S_n is primitive recursive.

Now (4) translates under the arithmetization into

(6a) $$(Ey)R(x_1, \cdots, x_n, y) \equiv (Ey)S_n(f, x_1, \cdots, x_n, y)$$

with f as the Gödel number of the system of equations F. The formula

(7a) $$(y)R(x_1, \cdots, x_n, y) \equiv (y)\overline{S}_n(g, x_1, \cdots, x_n, y)$$

is obtained likewise from (5), after changing the notation so that R is interchanged with $\overline{R}$.

In stating these results for reference, we shall go over from S_n to a new predicate T_n, which entails no present disadvantage and proves to be of convenience in some further investigations(7). The predicate T_n is defined from S_n as follows.

$$T_n(z, x_1, \cdots, x_n, y)\colon\ S_n(z, x_1, \cdots, x_n, y)\ \&\ (t)\left[t < y \rightarrow \overline{S}_n(z, x_1, \cdots, x_n, t)\right].$$

By a theorem of Gödel(8), the primitive recursiveness of T_n follows from that of S_n. The formulas (6) and (7) in the theorem follow from (6a) and (7a) by the definition of T_n in terms of S_n.

THEOREM I. *Given a general recursive predicate* $R(x_1, \cdots, x_n, y)$, *there are numbers* f *and* g *such that*

(6) $$(Ey)R(x_1, \cdots, x_n, y) \equiv (Ey)T_n(f, x_1, \cdots, x_n, y),$$

(7) $$(y)R(x_1, \cdots, x_n, y) \equiv (y)\overline{T}_n(g, x_1, \cdots, x_n, y).$$

Now $(Ey)T_n(z, x_1, \cdots, x_n, y)$ is a fixed predicate of the form $(Ey)R(z, x_1, \cdots, x_n, y)$ where R is general recursive (in fact, as it happens, primitive recursive). By the theorem, if we take successively $z = 0, 1, 2, \cdots$, we obtain an enumeration (with repetitions) of all predicates of the form $(Ey)R(x_1, \cdots, x_n, y)$ where R is general recursive(9). Likewise, the theorem gives us a fixed predicate of the form $(y)R(z, x_1, \cdots, x_n, y)$ where R is general recursive which enumerates all predicates of the form $(y)R(x_1, \cdots, x_n, y)$

(6) Gödel [1].

(7) A revision, April 13, 1942.

(8) Gödel [1, IV].

(9) This result entered partly into the last theorem of Kleene [2], but the advantage of using it at an earlier stage was overlooked. In anticipation, we may remark that XI–XVI of that paper are essentially special cases of Theorem II below (with now a constructive proof for XVI).

where R is general recursive. These enumerations form the basis for the application of Cantor's diagonal method in the next section.

5. **The general theorem.** By a familiar rule of classical logic, in each of the following pairs of propositions (with a fixed R for a given pair),

$$\begin{array}{cccc} (Ex)R(x) & (x)(Ey)R(x, y) & (Ex)(y)(Ez)R(x, y, z) & \cdots \\ (x)\overline{R}(x) & (Ex)(y)\overline{R}(x, y) & (x)(Ey)(z)\overline{R}(x, y, z) & \cdots, \end{array}$$

either member is equivalent to the negation of the other. Hence we may assert non-equivalence between the members of the pair. This argument is not good in the intuitionistic logic. However, the non-equivalence for the case of one quantifier,

$$(8) \qquad (Ex)R(x) \not\equiv (x)\overline{R}(x),$$

does hold good intuitionistically.

Consider the predicate form $(x)R(a, x)$ where R is general recursive. This gives a particular predicate of the variable a, whenever we specify the general recursive predicate $R(a, x)$ of two variables. In particular, $(x)\overline{T}_1(a, a, x)$ is a predicate of this form.

We shall show that this predicate is neither general recursive nor expressible in the form $(Ex)R(a, x)$ where R is general recursive.

For this purpose, suppose we have selected any particular general recursive $R(a, x)$, giving a particular predicate of the latter form. By (6), there is for this R a number f such that

$$(9) \qquad (Ex)R(a, x) \equiv (Ex)T_1(f, a, x).$$

Substituting the number f for the variable a,

$$(10) \qquad (Ex)R(f, x) \equiv (Ex)T_1(f, f, x).$$

By (8),

$$(11) \qquad (Ex)T_1(f, f, x) \not\equiv (x)\overline{T}_1(f, f, x).$$

Combining (10) and (11),

$$(12) \qquad (Ex)R(f, x) \not\equiv (x)\overline{T}_1(f, f, x).$$

This refutes, for $a=f$, the equivalence of $(Ex)R(a, x)$ to $(x)\overline{T}_1(a, a, x)$. Since this refutation can be effected, whatever general recursive R we chose, for some f depending on the R, the predicate $(x)\overline{T}_1(a, a, x)$ is not expressible in the form $(Ex)R(a, x)$ where R is general recursive.

A fortiori, $(x)\overline{T}_1(a, a, x)$ is not expressible in the form $R(a)$ where R is general recursive. For were it so expressed, we should then have it in the form $(Ex)R(a, x)$ where R is general recursive, by taking as $R(a, x)$ the predicate $R(a)$ & $x=x$.

This completes the proof of one case of the next theorem.

For another case, consider the predicate form $(Ex)R(a, x)$ where R is general recursive. We can show similarly, using (7) instead of (6), that the predicate $(Ex)T_1(a, a, x)$, which has this form, is neither general recursive nor expressible in the form $(x)R(a, x)$ where R is general recursive.

To illustrate the treatment of a case with more than one quantifier, consider the predicate form $(x)(Ey)(z)R(a, x, y, z)$ where R is general recursive. The predicate $(x)(Ey)(z)\overline{T}_3(a, a, x, y, z)$ has this form. Select any particular general recursive $R(a, x, y, z)$. By (6), for some f depending on this R,

$$(Ez)R(a, x, y, z) \equiv (Ez)T_3(f, a, x, y, z). \tag{13}$$

By corresponding quantifications of these equivalent predicates,

$$(Ex)(y)(Ez)R(a, x, y, z) \equiv (Ex)(y)(Ez)T_3(f, a, x, y, z). \tag{14}$$

Classically, we can complete the argument as before, showing that $(x)(Ey)(z)\overline{T}_3(a, a, x, y, z)$ is not expressible in any of the forms

$$R(a) \quad \begin{array}{lll} (Ex)R(a, x) & (x)(Ey)R(a, x, y) & (Ex)(y)(Ez)R(a, x, y, z) \\ (x)R(a, x) & (Ex)(y)R(a, x, y) & \end{array}$$

where the R for the form is general recursive.

To obtain an alternative phrasing of the theorem, in which it holds for all cases intuitionistically, we may omit in the classical proof the step which interchanges the two kinds of quantifiers under the operation of negation. We thus show that the predicates $\overline{(Ex)}T_1(a, a, x)$, $\overline{(x)}\overline{T}_1(a, a, x)$, $\overline{(Ex)}(y)(Ez)T_3(a, a, x, y, z)$, and so on, are neither expressible in the respective forms $(Ex)R(a, x)$, $(x)R(a, x)$, $(Ex)(y)(Ez)R(a, x, y, z)$, and so on, where R is general recursive, nor in any of the forms with fewer quantifiers.

THEOREM II. *Classically, and for the one-quantifier forms intuitionistically: To each of the forms*

$$R(a) \quad \begin{array}{lll} (Ex)R(a, x) & (x)(Ey)R(a, x, y) & (Ex)(y)(Ez)R(a, x, y, z) \cdots \\ (x)R(a, x) & (Ex)(y)R(a, x, y) & (x)(Ey)(z)R(a, x, y, z) \cdots \end{array}$$

where the R for each is general recursive, after the first, there is a predicate expressible in that form but not in the other form with the same number of quantifiers nor in any of the forms with fewer quantifiers.

Classically, and intuitionistically: To each of the forms, after the first, there is a predicate expressible in the negation of that form but not in that form itself nor in any of the forms with fewer quantifiers.

For simplicity, we have given the theorem for predicates of one variable a, but it holds:

Likewise, replacing the variable a throughout by n variables $a_1, \cdots, a_n$, for any fixed positive integer n.

By an *elementary predicate*, we shall mean one which is expressible in terms of general recursive predicates, the operations &, $\vee$, $^{-}$, $\rightarrow$, $\equiv$ of the propositional calculus, and quantifiers.

Suppose given an expression for a predicate in these terms. By the classical predicate calculus, we can transform the expression so that all quantifiers stand at the front. For each m, let $(x)_1, \cdots, (x)_m$ be a set of m primitive recursive functions of x which as a set ranges, with or without repetitions, over all m-tuples of natural numbers, as x ranges over all natural numbers (such sets of functions are known). The equivalences

$$(15) \qquad (Ex_1) \cdots (Ex_m)A(x_1, \cdots, x_m) \equiv (Ex)A((x)_1, \cdots, (x)_m),$$

$$(16) \qquad (x_1) \cdots (x_m)A(x_1, \cdots, x_m) \equiv (x)A((x)_1, \cdots, (x)_m)$$

enable us to eliminate consecutive occurrences of like quantifiers. These transformations leave as operand of the prefixed quantifiers a general recursive predicate of the free and bound variables. Hence, classically, the predicate forms listed in the theorem for a given n suffice for the expression of every elementary predicate of n variables.

The theorem then says that no finite sublist of the forms would suffice.

Classically, we are led to a classification of the elementary predicates according to the minimum numbers of quantifiers which would suffice for their expression in terms of general recursive predicates and quantifiers.

The analogy between the logical operations of existential and universal quantification and geometrical operations of projection and intersection, respectively, is well known(10). The possibility of a connection between present results and theories of Borel and Baire is suggested(11).

II. Primitive, general, and partial recursive predicates under quantification

6. **Partial recursive functions.** The author's definition of *partial recursive* function extends the Herbrand-Gödel definition of general recursive function to functions ϕ of n variables which need not be defined for all n-tuples of natural numbers as arguments, retaining the characteristic of that definition with respect to each n-tuple for which the function is defined(12). The partial recursive functions include the general recursive functions as those which are defined for all sets of arguments.

For a more complete description, take the definition of general recursive function which is given at the end of §2, and replace the four capitalized phrases by the following, respectively: *PARTIAL RECURSIVE*; PARTIAL RECURSIVE; AT MOST; THE NATURAL NUMBER x REPRE-

(10) In particular, it has been discussed by Tarski.

(11) This suggestion was made to the author by Gödel and by Ulam.

(12) Kleene [4].

SENTED BY THE NUMERAL x IF THAT NUMERAL EXISTS, AND IS OTHERWISE UNDEFINED.

In dealing with functions which may not be completely defined, we interpret the equation $\phi(x_1, \cdots, x_n) = \psi(x_1, \cdots, x_n)$ as the assertion that ϕ and ψ have the same value for $x_1, \cdots, x_n$ as arguments, taking it as undefined (non-significant) if either value is undefined. We write $\phi(x_1, \cdots, x_n) \simeq \psi(x_1, \cdots, x_n)$ to express the assertion that, if either of ϕ and ψ is defined for the arguments $x_1, \cdots, x_n$, the other is and the values are the same, and if either of ϕ and ψ is undefined for those arguments, the other is.

Similarly, in dealing with predicates which may not be completely defined, $P(x_1, \cdots, x_n) \equiv Q(x_1, \cdots, x_n)$ expresses equivalence of value, and is undefined if the value of either member is undefined; while $P(x_1, \cdots, x_n) \cong Q(x_1, \cdots, x_n)$ expresses that the definition of either implies mutual definition with equivalence, and the indefinition of either implies mutual indefinition.

A predicate $P(x_1, \cdots, x_n)$ not necessarily defined for all n-tuples of natural numbers as arguments is *partial recursive*, if there is a partial recursive function $\pi(x_1, \cdots, x_n)$ taking only 0 and 1 as values such that

$$P(x_1, \cdots, x_n) \cong \pi(x_1, \cdots, x_n) = 0; \tag{17}$$

in this case, π is called the *representing function* of P. (Or if we first introduce a *representing function* π of P, the value of which is to be 0, 1, or undefined according as the value of P is true, false, or undefined, then P is partial recursive if π is.)

In §§2, 3, we remarked the general recursiveness of Schemata (I)–(VI) with (VI) subjected to the condition (3); and we also considered Schema (VI) for the case that ρ is general recursive but (3) is not required to hold. The method of those sections applies equally well without the restrictions; in explanation of the schemata when the given functions may not be completely defined or (3) not hold for (VI), it will suffice here to remark that the conditions of definition for the functions introduced by the schemata may be inferred a posteriori from the metamathematical results.

THEOREM III. *The class of general recursive functions is closed under applications of Schemata* (I)–(VI) *with* (3) *holding for applications of* (VI).

The class of partial recursive functions is closed under applications of Schemata (I)–(VI).

COROLLARY. *Every function obtainable by applications of Schemata* (I)–(VI) *with* (3) *holding for applications of* (VI) *is general recursive.*

Every function obtainable by applications of Schemata (I)–(VI) *is partial recursive.*

7. **Normal form for recursive functions.** We shall pursue a little further

the method of §4 to obtain the converse of this result. Besides the metamathematical predicate $\mathfrak{S}_n$, we now require a metamathematical function as follows.

$\mathfrak{U}(\mathrm{Y})$: *the natural number x which the numeral* $\mathbf{x}$ *represents, in case* Y *is a formal deduction of an equation of the form* $\mathrm{t}=\mathbf{x}$, *where* $\mathbf{x}$ *is a numeral and* t *is any term; and* 0, *otherwise.*

According to the definition of general recursive function, if ϕ is a general recursive function of n variables, there is a system E of equations such that

$$(18) \qquad (x_1) \cdots (x_n)(E\mathrm{Y})\mathfrak{S}_n(\mathrm{E}, x_1, \cdots, x_n, \mathrm{Y}),$$

$$(19) \qquad (x_1) \cdots (x_n)(\mathrm{Y})[\mathfrak{S}_n(\mathrm{E}, x_1, \cdots, x_n, \mathrm{Y}) \rightarrow \mathfrak{U}(\mathrm{Y}) = \phi(x_1, \cdots, x_n)];$$

and the function $\phi(x_1, \cdots, x_n)$ can be expressed in terms of E thus

$$(20) \qquad \phi(x_1, \cdots, x_n) = \mathfrak{U}(\mu \mathrm{Y}\mathfrak{S}_n(\mathrm{E}, x_1, \cdots, x_n, \mathrm{Y})),$$

if we understand the formal objects to be enumerated in some order, so that the operator μ can be applied with respect to the metamathematical variable Y; we may take the order to be that of the corresponding Gödel numbers.

If ϕ is a partial recursive function of n variables, instead of asserting (18), we can write

$$(E\mathrm{Y})\mathfrak{S}_n(\mathrm{E}, x_1, \cdots, x_n, \mathrm{Y})$$

as the condition on $x_1, \cdots, x_n$ that the function be defined for $x_1, \cdots, x_n$ as arguments; we have (19), taking the implication to be true whenever the first member is false, irrespective of the status of the second member; and our convention calls for rewriting (20) thus,

$$(21) \qquad \phi(x_1, \cdots, x_n) \simeq \mathfrak{U}(\mu \mathrm{Y}\mathfrak{S}_n(\mathrm{E}, x_1, \cdots, x_n, \mathrm{Y})),$$

in order that it be true (and not sometimes undefined) for all values of $x_1, \cdots, x_n$.

By the Gödel numbering already considered, the metamathematical function $\mathfrak{U}(\mathrm{Y})$ is carried into a number-theoretic function $U(y)$, the definition of which we complete by taking the value to be 0 for any y not correlated to a formal object. If the Gödel numbering was suitably chosen, U as well as S_n is primitive recursive.

Now (20), (18) and (19) in terms of $\mathfrak{S}_n$ and $\mathfrak{U}$ are carried into formulas of like form in terms of S_n and U. On passing over from S_n to T_n, we then have the (22), (23) and (24) of the theorem[13]. The part of the theorem which refers to a partial recursive function is obtained similarly.

THEOREM IV. *Given a general recursive function* $\phi(x_1, \cdots, x_n)$, *there is a*

---

[13] Kleene [2, IV], with some changes in the formulation. The present S_n corresponds to the former T_n, using the Gödel numbering of proofs instead of the enumeration of provable equations.

number e such that

(22) $$\phi(x_1, \cdots, x_n) = U(\mu y T_n(e, x_1, \cdots, x_n, y)),$$

(23) $$(x_1) \cdots (x_n)(Ey)T_n(e, x_1, \cdots, x_n, y),$$

(24) $$(x_1) \cdots (x_n)(y)[T_n(e, x_1, \cdots, x_n, y) \rightarrow U(y) = \phi(x_1, \cdots, x_n)].$$

Given a partial recursive function $\phi(x_1, \cdots, x_n)$, there is a number e such that

(25) $$\phi(x_1, \cdots, x_n) \simeq U(\mu y T_n(e, x_1, \cdots, x_n, y)),$$

where

$$(Ey)T_n(e, x_1, \cdots, x_n, y)$$

is the condition of definition of the function, and (24) *holds.*

Thus any general recursive function (any partial recursive function) is expressible in the form $\psi(\mu y R(x_1, \cdots, x_n, y))$ with (2) holding (in the form $\psi(\mu y R(x_1, \cdots, x_n, y)))$ where ψ and R are primitive recursive. Hence:

COROLLARY. *Every general recursive function is obtainable by applications of Schemata* (I)–(VI) *with* (3) *holding for applications of* (VI).

Every partial recursive function is obtainable by applications of Schemata (I)–(VI).

Formula (25) contains the substance of the theorem. For it implies the condition of definition of the function; and, in the case that $\phi(x_1, \cdots, x_n)$ is defined for all sets of arguments, it gives (22) and (23). Moreover by the definition of T_n in terms of S_n, it implies (24).

We say that *e defines ϕ recursively*, or *e* is a *Gödel number* of ϕ, if (25) holds[14], in which case *e* has all the properties in relation to ϕ which are specified in the theorem.

It is here that the advantage of using T_n instead of S_n appears. A number *e* which satisfies $\phi(x_1, \cdots, x_n) \simeq U(\mu y S_n(e, x_1, \cdots, x_n, y))$ (which is equivalent to (25)) does not necessarily satisfy $(x_1) \cdots (x_n)(y)[S_n(e, x_1, \cdots, x_n, y) \rightarrow U(y) = \phi(x_1, \cdots, x_n)]$. While we could get around the difficulty by imposing the latter as an additional condition on the Gödel numbers, it is more convenient simply to use T_n instead of S_n. (On the basis of Theorem III and the results which we had in terms of S_n before passing over to T_n, one can set up a primitive recursive function V such that, if *e* satisfies (25), then $V(e)$ has all the properties in terms of S_n.)

The numbers *f* and *g* for Theorem I can be described now as any numbers which define recursively the partial recursive functions $\mu y R(x_1, \cdots, x_n, y)$ and $\mu y \overline{R}(x_1, \cdots, x_n, y)$, respectively.

---

(14) Kleene [2, Definition 2c, p. 738] and [4, top p. 153]. We have now also the changes in the formulation of Theorem IV.

8. **Consistency.** Let us review the arguments used in proof of Theorems I and III. For rigor, these have to be put in metamathematical form. Let E be the system of equations associated with a series of applications of Schemata (I)–(VI). We shall review only the case that no given function symbols occur in E.

In general, we easily establish that, for each of certain sets $x_1, \cdots, x_n$ of natural numbers, an equation of the form $\mathrm{f}(\mathbf{x}_1, \cdots, \mathbf{x}_n)=\mathbf{x}$, as described in the definitions of general and partial recursive function, is derivable from E by R1 and R2. In particular, if we are proving that E defines a general recursive function, we must show this for all $x_1, \cdots, x_n$; if we have a prior interpretation of the schemata applications as definition of a (partial or complete) function $\phi(x_1, \cdots, x_n)$, or require that E define a $\phi(x_1, \cdots, x_n)$ already known to us in some other manner, we must show this for all $x_1, \cdots, x_n$ belonging to the range of definition of ϕ, and also show that the $\mathbf{x}$ in the equation is the numeral representing the value of ϕ for $x_1, \cdots, x_n$ as arguments. This property of the equations E and rules R1 and R2, the precise formulation of which depends on the circumstances, we call the "completeness property." (When we wish merely to show that E defines a partial recursive function, the function to be determined a posteriori from E, no completeness property is required.)

The second part of the discussion consists in showing that an equation of the described form $\mathrm{f}(\mathbf{x}_1, \cdots, \mathbf{x}_n)=\mathbf{x}$ is derivable from E for at most one numeral $\mathbf{x}$; or if we have already established completeness in one of the above senses, that the equations $\mathrm{f}(\mathbf{x}_1, \cdots, \mathbf{x}_n)=\mathbf{x}$ referred to in the discussion of completeness, for various $x_1, \cdots, x_n$, are the only equations of that form which are derivable from E by R1 and R2. This we call the "consistency property."

As we indicated in §2, it suffices to handle each of the schemata in turn, assuming equations for use with R2 which give the values of the given functions. The argument for consistency which we sketched in §3 for Schema (VI) applies as well to the other schemata. For Schema (IV) there is indeed a choice in the order in which the values of the several χ's are introduced, but it is without effect on the final result.

This very easy consistency proof was gained by restricting the replacement rule so that replacement is only performable on the right member of an equation, a part $\mathrm{f}(\mathbf{x}_1, \cdots, \mathbf{x}_n)$ where f is a function symbol and $\mathbf{x}_1, \cdots, \mathbf{x}_n$ are numerals being replaced by a numeral $\mathbf{x}$. This eliminates the possibility of deriving an equation of the form $\mathrm{g}(\mathbf{y}_1, \cdots, \mathbf{y}_m)=\mathbf{y}$, where g is a fixed function symbol, $\mathbf{y}_1, \cdots, \mathbf{y}_m$ are fixed numerals, and $\mathbf{y}$ is any numeral, along essentially different paths within the system, and therewith the possibility that such an equation should be derivable for different $\mathbf{y}$'s.

In some previous versions of the theories of general and partial recursive functions, the replacement rule was not thus restricted. The consistency proof

which we gave in the version with the unrestricted replacement rule was based on the notion of verifiability of an equation(15). This notion makes presupposition of the values of the functions, and for the theory of partial recursive functions also of the determinateness whether or not the values are defined. In the latter case, it is not finitary. To give a constructive consistency proof for the theory of partial recursive functions with the stronger replacement rule seems to require the type of argument used in the Church-Rosser consistency proof for λ-conversion(16), and in the Ackermann-von Neumann consistency proof for a certain part of number theory in terms of the Hilbert ϵ-symbol(17).

It is easily shown, by using the method of proof of Theorem IV to obtain the same normal form with the stronger replacement rule, that every function partial recursive under the stronger replacement rule is such under the weaker.

Thus we find the curious fact that the main difficulty in showing the equivalence of the two notions of recursiveness comes in showing that the stronger rule suffices to define as many functions as the weaker. This is because the consistency of a stronger formalism is involved. The consistency of that formalism is of interest on its own account, but is extraneous for the theory of recursive definition, including the applications corresponding to those of Church in terms of the λ-notation which presuppose the complicated Church-Rosser consistency proof. All that is required for the theory of recursive definition is some consistent formalism sufficient for the derivation of the equations giving the values of the functions.

To this discussion we may add several supplementary remarks. We might in practice have a system E of equations and a method for deriving from E by R1 and the strong replacement rule, for all and only the n-tuples of a certain set, an equation of the form $\mathrm{f}(\mathbf{x}_1, \cdots, \mathbf{x}_n) = \mathbf{x}$ with a determinate $\mathbf{x}$, but lack the knowledge that unlimited use of the two rules could not lead to other such equations. In this situation, a function is defined intuitively for the n-tuples of the set, and undefined off the set. If we can characterize metamathematically our method of applying the two rules, we shall obtain a limited formalism known to be consistent, and the method used in establishing Theorem IV can then be applied to obtain equations defining the function recursively with the weak replacement rule.

For some types of equations which define a function recursively with the strong replacement rule (consistency being known), a more direct method may be available for obtaining a system defining the function recursively with the weak replacement rule. For example, consider (in informal language) the equation $\phi(\psi(x)) = \chi(x)$. To use this in deriving equations giving values of ϕ, we need to introduce values of ψ by replacement on the left. After

(15) Kleene [2, p. 731] and [4, §2, the bracketed portion of the fifth paragraph].
(16) Church and Rosser [1].
(17) Hilbert and Bernays [1, §2, part 4, pp. 93–130, and Supplement II, pp. 396–400].

expressing the equation in the form $\phi(y) = (\mu w[\psi((w)_1) = y \,\&\, \chi((w)_1) = (w)_2])_2$, and separating the latter into a series of equations without the μ-symbol by the method which the theory of the schemata affords, replacement will be required only on the right. This device is applicable to any equation of which the left member has the form $\mathrm{f}(\mathrm{g}_1(\mathrm{x}_1, \cdots, \mathrm{x}_n), \cdots, \mathrm{g}_m(\mathrm{x}_1, \cdots, \mathrm{x}_n))$.

The precise form of the restriction which is used to weaken the replacement rule is somewhat arbitrary, so long as it accomplishes its purpose of channelling the deductions of equations giving the values of the functions. The restriction as it was stated in the early Gödel version is now simplified, since we need to consider only equations having the forms appearing in the six schemata. Gödel provided for equations the left members of which could have the form $\mathrm{f}(\mathrm{g}_1(\mathrm{x}_1, \cdots, \mathrm{x}_n), \cdots, \mathrm{g}_m(\mathrm{x}_1, \cdots, \mathrm{x}_n))$ where f is the principal function symbol and $\mathrm{g}_1, \cdots, \mathrm{g}_m$ are given function symbols, and therefore allowed replacement on the left in the case of the g's.

9. **Predicates expressible in both one-quantifier forms.** By Theorem IV, for any general recursive predicate $P(x_1, \cdots, x_n)$,

$$(26) \qquad P(x_1, \cdots, x_n) \equiv (Ey)[T_n(e, x_1, \cdots, x_n, y) \,\&\, U(y) = 0],$$

$$(27) \qquad P(x_1, \cdots, x_n) \equiv (y)[T_n(e, x_1, \cdots, x_n, y) \rightarrow U(y) = 0],$$

where e is any Gödel number of the representing function of P.

Conversely, suppose that for a predicate P both $P(x_1, \cdots, x_n) \equiv (Ey)R(x_1, \cdots, x_n, y)$ and $P(x_1, \cdots, x_n) \equiv (y)S(x_1, \cdots, x_n, y)$ where R and S are general recursive. From the second of these equivalences, under classical interpretations, $\overline{P}(x_1, \cdots, x_n) \equiv (Ey)\overline{S}(x_1, \cdots, x_n, y)$. By the classical law of the excluded middle, $(Ey)[R(x_1, \cdots, x_n, y) \vee \overline{S}(x_1, \cdots, x_n, y)]$. Therefore

$$(28) \quad P(x_1, \cdots, x_n) \equiv R(x_1, \cdots, x_n, \mu y[R(x_1, \cdots, x_n, y) \vee \overline{S}(x_1, \cdots, x_n, y)]),$$

where the second member is general recursive by Theorem III.

THEOREM V. *Every general recursive predicate $P(x_1, \cdots, x_n)$ is expressible in both of the forms $(Ey)R(x_1, \cdots, x_n, y)$ and $(y)R(x_1, \cdots, x_n, y)$ where the R for each is primitive recursive. Under classical interpretations, conversely, every predicate expressible in both of these forms where the R for each is general recursive is general recursive.*

Now consider any predicate expressible in one of the forms of Theorem II after the first. According as the innermost quantifier in this form is existential or universal, we can apply (26) or (27), and then absorb the extra quantifier by (15) or (16), respectively, to obtain the original form but with a primitive recursive R. For example,

$$\begin{aligned}(x)(Ey)R(a, x, y) &\equiv (x)(Ey_1)(Ey_2)[T_3(e, a, x, y_1, y_2) \,\&\, U(y_2) = 0] \\ &\equiv (x)(Ey)[T_3(e, a, x, (y)_1, (y)_2) \,\&\, U((y)_2) = 0].\end{aligned}$$

COROLLARY. *The class of predicates expressible in a given one of the forms of Theorem* II *after the first (for a given n variables) is the same whether a primitive recursive or a general recursive R be allowed.*

This generalizes the observation of Rosser that a class enumerable by a general recursive function is also enumerable by a primitive recursive function([18]).

The formulas for the one-quantifier cases are

$$(29)\quad (Ey)R(x_1, \cdots, x_n, y) \equiv (Ey)[T_{n+1}(e, x_1, \cdots, x_n, (y)_1, (y)_2) \;\&\; U((y)_2) = 0],$$

$$(30)\quad (y)R(x_1, \cdots, x_n, y) \equiv (y)[T_{n+1}(e, x_1, \cdots, x_n, (y)_1, (y)_2) \rightarrow U((y)_2) = 0],$$

where e is any Gödel number of the representing function of R. These afford a new proof of the enumeration theorem of §4, with new enumerating predicates, and thence a new proof of Theorem II.

10. **Partial recursive predicates.** Let $P(x_1, \cdots, x_n)$ be a predicate which may not be defined for all n-tuples of natural numbers as arguments. By a *completion* of P we understand a predicate Q such that, if $P(x_1, \cdots, x_n)$ is defined, then $Q(x_1, \cdots, x_n)$ is defined and has the same value, and if $P(x_1, \cdots, x_n)$ is undefined, then $Q(x_1, \cdots, x_n)$ is defined. In particular, the completion $P^+(x_1, \cdots, x_n)$ which is false when $P(x_1, \cdots, x_n)$ is undefined, and the completion $P^-(x_1, \cdots, x_n)$ which is true when $P(x_1, \cdots, x_n)$ is undefined, we call the *positive completion* and *negative completion* of $P(x_1, \cdots, x_n)$, respectively. (In P and P^+, the "positive parts" coincide; in P and P^-, the "negative parts" coincide.)

If $P(x_1, \cdots, x_n)$ is a partial recursive predicate, then by Theorem IV,

$$(31)\quad P^+(x_1, \cdots, x_n) \equiv (Ey)[T_n(e, x_1, \cdots, x_n, y) \;\&\; U(y) = 0],$$

$$(32)\quad P^-(x_1, \cdots, x_n) \equiv (y)[T_n(e, x_1, \cdots, x_n, y) \rightarrow U(y) = 0],$$

where e is any Gödel number of the representing function of P.

Conversely, if $R(x_1, \cdots, x_n, y)$ is any general recursive predicate, then by Theorem III,

$$(33)\quad (Ey)R(x_1, \cdots, x_n, y) \equiv \mu y R(x_1, \cdots, x_n, y) =^+ \mu y R(x_1, \cdots, x_n, y),$$

$$(34)\quad (y)R(x_1, \cdots, x_n, y) \equiv \mu y \overline{R}(x_1, \cdots, x_n, y) \neq^- \mu y \overline{R}(x_1, \cdots, x_n, y).$$

THEOREM VI. *The positive completion $P^+(x_1, \cdots, x_n)$ of a partial recursive predicate $P(x_1, \cdots, x_n)$ is expressible in the form $(Ey)R(x_1, \cdots, x_n, y)$ where R is primitive recursive; and conversely, any predicate expressible in the form $(Ey)R(x_1, \cdots, x_n, y)$ where R is general recursive is the positive completion $P^+(x_1, \cdots, x_n)$ of a partial recursive predicate $P(x_1, \cdots, x_n)$.*

([18]) Rosser [1, Lemma I, Corollary I, p. 88].

Dually, for negative completions $P^-(x_1, \cdots, x_n)$ *and the predicate form* $(y)R(x_1, \cdots, x_n, y)$.

It follows that, for the predicate forms of Theorem II which have an existential quantifier (universal quantifier) innermost, we may, without altering the class of predicates expressible in that form, take R to be the positive completion (negative completion) of a partial recursive predicate.

Let us abbreviate $U(\mu y T_n(z, x_1, \cdots, x_n, y))$ as $\Phi_n(z, x_1, \cdots, x_n)$[19]. Then Φ_n is a fixed partial recursive function of $n+1$ variables, from which any partial recursive function ϕ of n variables can be obtained thus (rewriting (25)),

$$\phi(x_1, \cdots, x_n) \simeq \Phi_n(e, x_1, \cdots, x_n) \tag{35}$$

where e is any Gödel number of ϕ. Since for a constant z, $\Phi_n(z, x_1, \cdots, x_n)$ is always a partial recursive function of the remaining n variables, $\Phi_n(z, x_1, \cdots, x_n)$ therefore gives for $z=0, 1, 2, \cdots$ an enumeration (with repetitions) of the partial recursive functions of n variables. It follows that $\Phi_n(z, x_1, \cdots, x_n)=0$ is a partial recursive predicate of $n+1$ variables which enumerates (with repetitions) the partial recursive predicates of n variables.

This, seen in the light of Theorem VI, has as consequence the enumeration theorem of §2 (with other enumerating predicates), and thence by Cantor's diagonal method Theorem II.

Elsewhere, the enumeration theorem for partial recursive functions gave by Cantor's diagonal method what may be called the fundamental theorem for proofs of recursive definability[20].

This fundamental theorem, and the existence of partial recursive functions and predicates, no completions of which are general recursive[21], are what occasioned the introduction of the notion of a partial recursive function.

III. Incompleteness theorems in the foundations of number theory

11. **Introductory remarks.** We entertain various propositions about natural numbers. These propositions have meaning, independently of or prior to the consideration of formal postulates and rules of proof. We pose the problem of systematizing our knowledge about these propositions into a theory of some kind. For certain definitions of our objectives in constructing the theory, and certain classes of propositions, we shall be able to reach definite answers concerning the possibility of constructing the theory.

The naïve informal approach which we are adopting may be contrasted

---

(19) Using the notation of Kleene [4, bottom p. 152], but with the changes in the formulation of Theorem IV.

(20) Kleene [4, the last result in §2].

(21) Kleene [4, Footnote 3].

with that form of the postulational approach which consists in first listing formal postulates, which are then said to define the content of the theory based on them. In the case of number theory, the formal approach cannot render entirely dispensable an intuitive understanding of propositions of the kind which we commonly interpret the theory to be about. For the explicit statement of the postulates and characterization of the manner in which they are to determine the theory belong to a metatheory on another level of discourse; and the ultimate metatheory must be an intuitive mathematics unregulated by explicit postulates, and having the essential character of number theory.

Of course the informality of our investigation does not preclude the enumeration, from another level, of postulates which would suffice to describe it. Indeed, such regulation may perhaps be considered necessary from an intuitive standpoint for that part of it which belongs to the context of classical mathematics.

The propositions about natural numbers which we shall consider will contain parameters. We shall thus have infinitely many propositions of a given form, according to the natural numbers taken as values by the parameters. In other words, we have predicates, for which these parameters are the independent variables. Generally, in a theory, a number of predicates are dealt with simultaneously; but for our investigations it will suffice to consider a theory with respect to some one predicate without reference to other predicates which might be present. Usually, we shall write a one-variable predicate $P(a)$, though the discussion applies equally well to a predicate $P(a_1, \cdots, a_n)$ of n variables.

12. **Algorithmic theories.** As one choice of the objective, we can ask that the theory should give us an effective means for deciding, for any given one of the propositions which are taken as values of the predicate, whether that proposition is true or false. Examples of predicates for which a theoretical conquest of this kind has been obtained are: *a is divisible by b* (that is, in symbols, $(Ex)[a=bx]$), *$ax+by=c$ is solvable for x and y* (that is, $(Ex)(Ey)[ax+by=c]$). We shall call this kind of theory for a predicate a ***complete algorithmic theory*** for the predicate.

Let us examine the notion of this kind of theory more closely. In setting up a complete algorithmic theory, what we do is to describe a procedure, performable for each set of values of the independent variables, which procedure necessarily terminates and in such manner that from the outcome we can read a definite answer, "Yes" or "No," to the question, "Is the predicate value true?"

We can express this by saying that we set up a second predicate: ***the procedure terminates in such a way as to give the affirmative answer.*** The second predicate has the same independent variables as the first, is equivalent to the first, and the determinability of the truth or falsity of its values is guaranteed.

This last property of the second predicate we designate as the property of being *effectively decidable*.

Of course the original predicate becomes effectively decidable, in a derivative sense, as soon as we have its equivalence to the second; extensionally, the two are the same. But while our terminology is ordinarily extensional, at this point the essential matter can be emphasized by using the intensional language. The reader may if he wishes write in more explicit statements referring to the (generally) differing objects or processes with which the two predicates are concerned.

Now, the recognition that we are dealing with a well defined process which for each set of values of the independent variables surely terminates so as to afford a definite answer, "Yes" or "No," to a certain question about the manner of termination, in other words, the recognition of effective decidability in a predicate, is a subjective affair. Likewise, the recognition of what may be called *effective calculability* in a function. We may assume, to begin with, an intuitive ability to recognize various individual instances of these notions. In particular, we do recognize the general recursive functions as being effectively calculable, and hence recognize the general recursive predicates as being effectively decidable.

Conversely, as a heuristic principle, such functions (predicates) as have been recognized as being effectively calculable (effectively decidable), and for which the question has been investigated, have turned out always to be general recursive, or, in the intensional language, equivalent to general recursive functions (general recursive predicates). This heuristic fact, as well as certain reflections on the nature of symbolic algorithmic processes, led Church to state the following thesis([22]). The same thesis is implicit in Turing's description of computing machines([23]).

THESIS I. *Every effectively calculable function (effectively decidable predicate) is general recursive.*

Since a precise mathematical definition of the term effectively calculable (effectively decidable) has been wanting, we can take this thesis, together with the principle already accepted to which it is converse, as a definition of it for the purpose of developing a mathematical theory about the term. To the extent that we have already an intuitive notion of effective calculability (effective decidability), the thesis has the character of an hypothesis—a point emphasized by Post and by Church([24]). If we consider the thesis and its converse as definition, then the hypothesis is an hypothesis about the application of the mathematical theory developed from the definition. For the acceptance of the hypothesis, there are, as we have suggested, quite compelling grounds.

([22]) Church [1].
([23]) Turing [1].
([24]) Post [1, p. 105], and Church [2].

A full account of these is outside the scope of the present paper(25). We are here concerned rather to present the consequences.

In the intensional language, to give a complete algorithmic theory for a predicate $P(a)$ now means to find an equivalent effectively decidable predicate $Q(a)$. It would suffice that $Q(a)$ be given as a general recursive predicate; and by Thesis I, if $Q(a)$ is not so given, then at least there is a general recursive predicate $R(a)$ equivalent to $Q(a)$ and hence to $P(a)$. Thus to give a complete algorithmic theory for $P(a)$ means to find an equivalent general recursive predicate $R(a)$, or more briefly, to express $P(a)$ in the form $R(a)$ where R is general recursive. This predicate form is the one listed first in Theorem II; and Theorem II gives to each of the other forms a predicate not expressible in that form. Thus, while under our interpretations there is a complete algorithmic theory for each predicate of the form $R(a)$ where R is general recursive, to each of the other forms there is a predicate for which no such theory is possible. We state this in the following theorem, using the particular examples for the one-quantifier forms which were exhibited in the proof of Theorem II.

THEOREM VII. *There exists no complete algorithmic theory for either of the predicates* $(Ex)T_1(a, a, x)$ *and* $(x)\overline{T}_1(a, a, x)$.

Of course, once the definition of effective decidability is granted, which affords an enumeration of the effectively decidable predicates, Cantor's methods immediately give other predicates. This theorem, as additional content, shows the elementary forms which suffice to express such predicates.

Abstracting from the particular examples used here, the theorem is Church's theorem on the existence of an unsolvable problem of elementary number theory, and the corresponding theorem of Turing in terms of his machine concept(26). The unsolvability is in the sense that the construction called for by the problem formulation, which amounts to that of a recursive R with a certain property, is impossible. The theorem itself constitutes solution in a negative sense.

13. **Formal deductive theories.** A second possibility for giving theoretic cohesion to the totality of true propositions taken as values of a predicate $P(a)$ is that offered by the postulational or deductive method. We should like all and only those of the predicate values which are true to be deducible from given axioms by given rules of inference. To make the axioms and principles of inference quite explicit, according to modern standards of rigor, we shall suppose them constituted into a formal system (symbolic logic), in which the propositions taken as values of the predicate are expressible. Those and only those of the formulas expressing the true instances of the predicate

(25) For a resume, see Kleene [4, Footnote 2], where further references are given.
(26) Turing [1, §8].

should be provable. We call this kind of theory for a predicate $P(a)$ a *complete formal deductive theory* for the predicate.

This type of theory should of course not be confused with incompletely formalized axiomatic theories, such as the theory of natural numbers itself as based on Peano's axioms.

It is convenient in discussing a formal system to name collectively as the "postulates" the rules describing the formal axioms and the rules of inference.

Let us now examine more closely the concept of provability in a stated formal system. If the formalization does accomplish its purpose of making matters explicit, we should be able effectively to recognize each step of a formal proof as an application of a postulate of the system. Furthermore, if the system is to constitute a theory for the predicate $P(a)$, we should be able effectively to recognize, to each natural number a, a certain formula of the system which is taken as expressing the proposition $P(a)$. Together, these conditions imply that we should be able, given any sequence of formulas which might be submitted as a proof of $P(a)$ for a given a, to check it, thus determining effectively whether it is actually such or not.

Let us introduce a designation for the metamathematical predicate with which we deal in making this check, for a given formal system and predicate $P(a)$.

$\mathfrak{R}(a, \mathrm{X})$: X *is a proof in the formal system of the formula expressing the proposition* $P(a)$.

Then the concept of provability in the system of the formula expressing $P(a)$, or briefly, the provability of $P(a)$, is expressible as $(E\mathrm{X})\mathfrak{R}(a, \mathrm{X})$.

As we have just argued, the predicate $\mathfrak{R}(a, \mathrm{X})$ should be an effectively decidable metamathematical predicate. Here the formal objects over which X ranges, if the notation of the system is explicit, should be given in some manner which affords an effective enumeration of them. Using the indices in this enumeration, or generally any effective Gödel numbering of the formal objects, the metamathematical predicate $\mathfrak{R}(a, \mathrm{X})$ will be carried into a number-theoretic predicate $R(a, x)$, taken as false for any x not correlated to a formal object, which should then also be effectively decidable. By Thesis I, the effective decidability of the latter implies its general recursiveness. We are thus led to state a second thesis.

THESIS II. *For any given formal system and given predicate* $P(a)$, *the predicate that* $P(a)$ *is provable is expressible in the form* $(Ex)R(a, x)$ *where* R *is general recursive.*

This thesis corresponds to the standpoint that the role of a formal deductive system for a predicate $P(a)$ is that of making explicit the notion of what constitutes a proof of $P(a)$ for a given a. If a proposed "formal system" for $P(a)$ does not do this, we should say that it is not a formal system in the

strict sense, or at least not one for $P(a)$. Taken this way, the thesis has a definitional character.

Presupposing, on the other hand, a prior conception of what constitutes a formal system for a given predicate in the strict sense, the thesis has the character of an hypothesis, to which we are led both heuristically and from Thesis I by general considerations.

Conversely, if a predicate of the form $(Ex)R(a, x)$ where R is general recursive is given, it is easily seen that we can always set up a formal system of the usual sort, with an explicit criterion of proof, in which all true instances of this predicate and only those are provable.

Using the thesis, and this converse, we can now say that to give a complete formal deductive theory for a predicate $P(a)$ means to find an equivalent predicate of the form $(Ex)R(a, x)$ where R is general recursive, or more briefly, to express the predicate in this form. By Theorem II, there are predicates of the other one-quantifier form, and of the forms with more quantifiers, not expressible in this form. Hence while there are complete formal deductive theories to each predicate of either of the forms $R(a)$ and $(Ex)R(a, x)$ where R is general recursive, to each of the other forms there is a predicate for which no such theory is possible. Specifically, using the one-quantifier example given in the proof of Theorem II:

THEOREM VIII. *There is no complete formal deductive theory for the predicate* $(x)\overline{T}_1(a, a, x)$.

This is the famous theorem of Gödel on formally undecidable propositions, in a generalized form. A proposition is formally undecidable in a given formal system if neither the formula expressing the proposition nor the formula expressing its negation is provable in the system. Gödel gave such a proposition for a certain formal system (by a method evidently applying to similar systems), subject to the assumptions of the consistency and ω-consistency of the system. Later Rosser gave another proposition, for which the latter assumption is dispensed with([27]).

In the present form of the theorem, we have a preassigned predicate $(x)\overline{T}_1(a, a, x)$ and a method which, to any formal system whatsoever for this predicate, gives a number f for which the following is the situation.

Suppose that the system meets the condition that the formula expressing the proposition $(x)\overline{T}_1(f, f, x)$ is provable only if that proposition is true. Then the proposition is true but the formula expressing it unprovable. This statement of results uses the interpretation of the formula, but if the system has certain ordinary deductive properties for the universal quantifier and recursive predicates, our condition on the system is guaranteed by the metamathematical one of consistency.

If the system contains also a formula expressing the negation of

([27]) Rosser [1].

$(x)\overline{T}_1(f, f, x)$, and if the system meets the further condition that this formula is provable only if true, then this formula cannot be provable, and we have a formally undecidable proposition. The further condition, if the system has ordinary deductive properties, is guaranteed by the metamathematical one of ω-consistency.

Moreover, we can incorporate Rosser's elimination of the hypothesis of ω-consistency into the present treatment. To do so, we replace the predicate $(Ex)R(a, x)$ for the application of Theorem II by $(Ex)[R(a, x) \ \& \ (y)[y<x \rightarrow \overline{S}(a, y)]]$ where $(Ex)S(a, y)$ is the predicate expressing the provability of the negation of $(x)\overline{T}_1(a, a, x)$. This changes the f for the system.

Thus we come out with the usual metamathematical results for a given formal system.

For the case that a formal system is sought which should not only prove the true instances of $P(a)$ but also refute the false ones, if the classical law of the excluded middle is applied to the propositions $P(a)$, then the Gödel theorem (Theorem VIII) comes under the Church theorem (Theorem VII). For had we completeness with respect both to $P(a)$ and to $\overline{P}(a)$, we could obtain a general recursive $R(a)$ equivalent to the given predicate by the method used in proving the second part of Theorem V. Informally, this amounts merely to the remark that we should have the algorithm for $P(a)$ which consists in searching through some list of the provable formulas until we encounter either the formula expressing $P(a)$ or the formula expressing $\overline{P}(a)$.

The connection between Gödel's theorem and the paradoxes has been much noted. The author gave a proof of Gödel's theorem along much the present lines but as a refinement of the Richard paradox rather than of the Epimenides([28]). That gave the undecidable propositions as values of a predicate of the more complicated form $(x)(Ey)R(a, x, y)$ where R is general recursive. The Epimenides paradox now appears as the more basic. Currently, Curry has noted the same phenomenon in connection with the Kleene-Rosser inconsistency theorem([29]).

14. **Discussion, incomplete theories.** In the present form of Gödel's theorem, several aspects are brought into the foreground which perhaps were not as clearly apparent in the original version.

Not merely, to any given formal system of the type considered, can a proposition be formulated with respect to which that system is incomplete, but all these propositions can be taken as values of a preassignable elementary predicate, with respect to which predicate therefore no system can be complete. This depends on the thesis giving a preassignable form to the concept of provability in a formal system.

([28]) Kleene [2, XIII].

([29]) Kleene and Rosser [1], Curry [2].

For the interpretation of the propositions we have required, as minimum, only the notions of effectively calculable predicates and of the quantifiers used constructively. It seems that lesser presuppositions, if one is to allow any mathematical infinite, are hardly conceivable.

Beyond that the system should fulfil the structural characteristic expressed in Thesis II, and should yield results correct under this modicum of interpretation, we have need of no reference whatsoever to its detailed constitution.

In particular, the nature of the intuitive evidence for the deductive processes which are formalized in the system plays no role.

Let us imagine an omniscient number theorist, whom we should expect, through his ability to see infinitely many facts at once, to be able to frame much stronger systems than any we could devise. Any correct system which he could reveal to us, telling us how it works without telling us why, would be equally subject to the Gödel incompleteness.

It is impossible to confine the intuitive mathematics of elementary propositions about integers to the extent that all the true theorems will follow from explicitly stated axioms by explicitly stated rules of inference, simply because the complexity of the predicates soon exceeds the limited form representing the concept of provability in a stated formal system.

We selected as the objective in constructing a formal deductive system that what constitutes proof should be made explicit in the sense that a proposed proof could be effectively checked, and either declared formally correct or declared formally incorrect.

Let us for the moment entertain a weaker conception of a formal system, under which, if we should happen to discover a correct proof of a proposition or be presented with one, then we could check it and recognize its formal correctness, but if we should have before us an alleged proof which is not correct, then we might not be able definitely to locate the formal fallacy. In other words, under this conception a system possesses a process for checking, which terminates in the affirmative case, but need not in the negative. Then the concept of provability would have the form $(Ex)P^+(a, x)$ where P^+ is the positive completion of a partial recursive predicate $P(a, x)$. By Theorem VI, $P^+(a, x)$ is expressible in the form $(Ey)R(a, x, y)$ where R is general recursive. Then the provability concept has the form $(Ex)(Ey)R(a, x, y)$, or by contraction of quantifiers $(Ex)R(a, (x)_1, (x)_2)$. This is of the form $(Ex)R(a, x)$ where R is general recursive. Thus the concept of provability has the usual form, and Gödel's theorem applies as before. If we take a new concept of proof based on $R(a, x)$, that is, if we redesignate the steps in the checking process as the formal proof steps, the concept of proof assumes the usual form.

We gave no attention, when we formulated the objectives both of an algorithmic and of a formal deductive theory, to the nature of the evidence for the correctness of the theory, or to various other practical considerations,

simply because the crude structural objectives suffice to entail the corresponding incompleteness theorems. In this connection, it may be of some interest to give the corresponding definitions, although these may not take into account all the desiderata, for the case of incomplete theories of the two sorts. We shall state these for predicates of n variables $a_1, \cdots, a_n$, as we could also have done for the case of the complete theories.

To give an *algorithmic theory* (not necessarily complete) for a predicate $P(a_1, \cdots, a_n)$ is to give a general recursive function $\pi(a_1, \cdots, a_n)$, taking only 0, 1, and 2 as values, such that

$$(36) \qquad \begin{cases} \pi(a_1, \cdots, a_n) = 0 \rightarrow P(a_1, \cdots, a_n) \\ \pi(a_1, \cdots, a_n) = 1 \rightarrow \overline{P}(a_1, \cdots, a_n). \end{cases}$$

The algorithm always terminates, but if $\pi(a_1, \cdots, a_n)$ has the value 2 we can draw no conclusion about $P(a_1, \cdots, a_n)$.

To give a *formal deductive theory* (not necessarily complete) for a predicate $P(a_1, \cdots, a_n)$ is to give a general recursive predicate $R(a_1, \cdots, a_n, x)$ such that

$$(37) \qquad (Ex)R(a_1, \cdots, a_n, x) \rightarrow P(a_1, \cdots, a_n).$$

In words, to give a formal deductive theory for a predicate $P(a_1, \cdots, a_n)$ is to find a sufficient condition for it of the form $(Ex)R(a_1, \cdots, a_n, x)$ where R is general recursive. Here, according to circumstances, the sufficiency may be established from a wider context, or it may be a matter of postulation (hypothesis), or of conviction (belief).

From the present standpoint, the setting up of this sufficient condition is the essential accomplishment in the establishment of a so-called metatheory (in the constructive sense) for the body of propositions taken as the values of a predicate. We note that this may be accomplished without necessarily going through the process of setting up a formal object language, from which R is obtainable by subsequent arithmetization, although as remarked above, we can always set up the object language, if we have the R by some other means.

In the view of the present writer, the interesting variations of formal technique recently considered by Curry have the above as their common feature with formalization of the more usual sort([30]). This is stated in our terminology, Curry's use of the terms "meta" and "recursive" being different. He gives examples of "formal systems," in connection with which he introduces some predicates by what he calls "recursive definitions," but what we should prefer to call "inductive definitions." This important type of definition, under suitable precise delimitation so that the individual clauses are constructive, can be shown to lead always to predicates expressible in the form $(Ex)R(a_1, \cdots, a_n, x)$ where R is recursive in our sense. Indeed, this fact

([30]) Curry [1].

can be recognized by substantially the method indicated above for the case of the inductive definition establishing the notion of provability for a formal system of the usual sort.

Conversely, given any predicate expressible in the form $(Ex)R(a_1, \cdots, a_n, x)$ where R is recursive, we can set up an inductive definition for it.

15. **Ordinal logics.** In ordinal logics, studied by Turing(31), the requirement of effectiveness for the steps of deduction is relaxed to allow dependence on a number (or λ-formula) which represents an ordinal in the Church-Kleene theory of constructive ordinals(32). A presumptive proof in an ordinal logic cannot in general be checked objectively, since the proof character depends on the number which occupies the role of a Church-Kleene representative of an ordinal actually being such, for which there is no effective criterion. Nevertheless it was hoped that ordinal logics could be used to give complete orderings (with repetitions) of the true propositions of certain forms into transfinite series, by means of the ordinals represented in the proofs, in such a way that the proving of a proposition in the ordinal logic (and therewith the determination of a position for it in the series) would somehow make it easier to recognize the truth of the proposition.

Turing obtained a number of interesting results, largely outside the scope of this article, but among them the following. There are ordinal logics which are complete for the theory of a predicate of the form $(x)(Ey)R(a, x, y)$ where R is general recursive; however, for the example of such a logic which is given, its use would afford no theoretic gain, since the recognition that the number which plays the role of ordinal representative in a proof of the logic is actually such comes to the same as the direct recognition of the truth of the proposition proved.

Now let us approach the topic by inquiring whether, and if so where, the property of being provable in a given ordinal logic is located in the scale of predicate forms of Theorem II. First, it turns out that the property of a number a of being the representative of an ordinal is expressible in the form $(x)(Ey)R(a, x, y)$ where R is recursive(33). Now we may use the definition of ordinal logic in terms of λ-conversion, or we may take the notion in general terms as described above, and state the thesis that for a given predicate $P(a)$ and given ordinal logic the provability of $P(a)$ is expressible in the form $(E\alpha)(Ex)R(a, \alpha, x)$ where α ranges over the ordinal representatives and R is general recursive. In either case, it then follows that the provability of $P(a)$ is expressible in the form $(Ex)(y)(Ez)R(a, x, y, z)$ where R is general recursive. Conversely, to any predicate of the latter form, we can find an ordinal logic

(31) Turing [2]. Turing gave a somewhat restricted definition of "ordinal logic" in terms of the theory of λ-conversion for predicates expressible in the form $(x)(Ey)R(a, x, y)$ where R is recursive.

(32) Church and Kleene [1], Church [2], Kleene [4].

(33) Kleene [5].

in the more general sense such that provability in the logic expresses the predicate. Hence there is a complete ordinal logic to each predicate of each of the forms

$$R(a) \quad \begin{matrix} (Ex)R(a, x) & (x)(Ey)R(a, x, y) & (Ex)(y)(Ez)R(a, x, y, z) \\ (x)R(a, x) & (Ex)(y)R(a, x, y) & \end{matrix}$$

where R is general recursive, but by Theorem II, classically there are predicates of the form $(x)(Ey)(z)R(a, x, y, z)$ and of each of the forms with more quantifiers, or classically and intuitionistically of the form $\overline{(Ex)}(y)(Ez)R(a, x, y, z)$ and of the negation of each of the forms with more quantifiers, for which no complete ordinal logic is possible. Specifically:

THEOREM IX. *There is no complete ordinal logic for the predicate* $\overline{(Ex)}(y)(Ez)T_3(a, a, x, y, z)$.

Ordinal logics form a class of examples of the systems of propositions which have recently come under discussion, in which more or less is retained of the ordering of propositions in deductive reasoning, but with an extension into the transfinite, or a sacrifice of constructiveness in individual steps. These may be called "non-constructive logics," in contrast to the formal deductive systems in the sense of §§13–14 which are "constructive logics." In general, the usefulness of a non-constructive logic may be considered to depend on the degree to which the statement of the non-constructive proof criterion is removed from the direct statement of the propositions.

Theorem IX is a "Gödel theorem" for the ordinal logics. The ordinal logics were at least conceived with somewhat of a constructive bias. Rosser has shown how Gödel theorems arise on going very far in the direction of non-constructiveness([34]), and Tarski has stated the Gödel argument for systems of sentences in general([35]). Incidental of Rosser's results for finite numbers of applications of the Hilbert "rule of infinite induction," also called "Carnap's rule," can easily be inferred from Theorem II, through the obvious correspondence of an application of this rule to a universal quantifier in the proof concept. However, the proof concepts for non-constructive logics soon outrun the scale of predicate forms of Theorem II. This appears to be the case even for the extension to protosyntactical definability given by Quine([36]). If one is going very far in the direction of non-constructiveness, and is not interested in considerations of the sort emphasized in §§12–14, there is no advantage in starting from the theory of recursive functions. But the more general results do not detract from the special significance which attaches to the Gödel theorems associated with provability criteria of the forms $R(a)$ and $(Ex)R(a, x)$

([34]) Rosser [2].
([35]) Tarski [2].
([36]) Quine [1].

where R is general recursive, that is, Church's theorem and Gödel's theorem, for which forms only it is true that a given proof is a finite object.

16. Constructive existence proofs. A proof of an existential proposition $(Ey)A(y)$ is acceptable to an intuitionist, only if in the course of the proof there is given a y such that $A(y)$ holds, or at least a method by which such a y could be constructed. Consider the case that $A(y)$ depends on other variables. Say that there is one of these, x, and rewrite the proposition as $(x)(Ey)A(x, y)$. The proposition asserts the existence of a y to each of the infinitely many values of x. In this case, the only way in which the constructivist demand could in general be met would be by giving the y as an effectively calculable function of x, that is, by giving the function. According to Thesis I, this function would have to be general recursive. Hence we propose the following thesis (and likewise for n variables $x_1, \cdots, x_n$):

THESIS III. *A proposition of the form $(x)(Ey)A(x, y)$ containing no free variables is provable constructively, only if there is a general recursive function $\phi(x)$ such that $(x)A(x, \phi(x))$.*

When such a ϕ exists, we shall say that $(x)(Ey)A(x, y)$ is *recursively fulfillable*(37).

This thesis expresses what seems to be demanded from the standpoint of the intuitionists. Whether such explicit rules of proof as they have stated do conform to the thesis is a further question which will be considered elsewhere(38). However, in its aspect as restriction on all intuitionistic existence proofs, the possibilities for which, as we know by Theorem VIII, transcend the limitations of any preassignable formal system, the thesis is more general than a metamathematical result concerning a given system.

We now examine the notion of recursive fulfillability as it applies to the values of a given predicate of the form $(x)(Ey)(z)R(a, x, y, z)$ where R is general recursive. Select any fixed value of a. Given a recursive ϕ which fulfils the corresponding proposition, by Theorem IV there is a number e such that $(x)(Ey)T_1(e, x, y)$ and $(x)(y)[T_1(e, x, y)\rightarrow(z)R(a, x, U(y), z)]$. Conversely, if such an e exists, the proposition is fulfilled by the general recursive function $U(\mu y T_1(e, x, y))$. Thus

$$(Ee)\{(x)(Ey)T_1(e, x, y) \ \& \ (x)(y)[T_1(e, x, y) \rightarrow (z)R(a, x, U(y), z)]\}$$

is a necessary and sufficient condition for recursive fulfillability. When the quantifiers are suitably brought to the front and contracted, this assumes the form $(Ex)(y)(Ez)R(a, x, y, z)$ with another general recursive R depending on the original R.

By Theorem II, classically, there is a predicate of the original form

(37) A further analysis of the implications of constructive provability is given in Kleene [6].

(38) Nelson [1].

$(x)(Ey)(z)R(a, x, y, z)$ which is not expressible in this form $(Ex)(y)(Ez)R(a, x, y, z)$, in which the condition of its recursive fulfillability is expressible.

Using the example of such a predicate given in the proof of Theorem II, we have then

$$(38) \qquad \{(x)(Ey)(z)\overline{T}_3(a, a, x, y, z) \text{ rec. fulf.}\} \equiv (Ex)(y)(Ez)R(a, x, y, z)$$

for a certain general recursive R. Substituting the number f of (14) for a in (14) and (38),

$$(39) \qquad (Ex)(y)(Ez)R(f, x, y, z) \equiv (Ex)(y)(Ez)T_3(f, f, x, y, z),$$

$$(40) \qquad \{(x)(Ey)(z)\overline{T}_3(f, f, x, y, z) \text{ rec. fulf.}\} \equiv (Ex)(y)(Ez)R(f, x, y, z).$$

By the definition of recursive fulfillability,

$$(41) \qquad \{(x)(Ey)(z)\overline{T}_3(f, f, x, y, z) \text{ rec. fulf.}\} \to (x)(Ey)(z)\overline{T}_3(f, f, x, y, z).$$

Suppose that $(x)(Ey)(z)\overline{T}_3(f, f, x, y, z)$ were recursively fulfillable. We could then conclude by (40) and (39), $(Ex)(y)(Ez)T_3(f, f, x, y, z)$, and by (41), $(x)(Ey)(z)\overline{T}_3(f, f, x, y, z)$. These results are incompatible. Therefore by reductio ad absurdum, $(x)(Ey)(z)\overline{T}_3(f, f, x, y, z)$ is not recursively fulfillable, and hence by Thesis III not constructively provable.

Now by (40) and (39), we have $\overline{(Ex)}(y)(Ez)T_3(f, f, x, y, z)$; and thence classically we can proceed to $(x)(Ey)(z)\overline{T}_3(f, f, x, y, z)$.

THEOREM X. *For a certain number f, the proposition $(x)(Ey)(z)\overline{T}_3(f, f, x, y, z)$ is true classically, but not constructively provable.*

Notice that we have here a fixed unprovable proposition for all constructive methods of reasoning, whereas in the preceding incompleteness theorems. we had only an infinite class of propositions, some of which must be unprovable in a given theory.

Intuitionistic number theory has been presented as a subsystem of the classical, so that the intuitionistic results hold classically, though many classical results are not asserted intuitionistically. The possibility now appears of extending intuitionistic number theory by incorporating Thesis III in the form

$$(x)(Ey)A(x, y) \to \{\text{for some general recursive } \phi,\ (x)A(x, \phi(x))\},$$

so that the two number theories should diverge, with the proposition of Theorem X true classically, and its negation true intuitionistically(39).

For the classical proof, an application of

$$\overline{(x)}A(x) \to (Ex)\overline{A}(x)$$

suffices as the sole non-intuitionistic step; therewith that law of logic would

(39) This is perhaps hinted in Church [1, first half of p. 363].

be refuted intuitionistically, for a certain A. Hitherto the intuitionistic refutations of laws of the classical predicate calculus have depended on the interpretation of the quantifiers in intuitionistic set theory(40).

The result of Theorem X, with another proposition as example, can be reached as follows. Consider the proposition,

$$(x)(Ey)\{[(Ez)T_1(x, x, z) \,\&\, y = 0] \vee [(z)\overline{T}_1(x, x, z) \,\&\, y = 1]\}.$$

This holds classically, by application of the law of the excluded middle in the form

$$(x)\{(Ez)A(x, z) \vee (z)\overline{A}(x, z)\},$$

or the form

$$(x)(A(x) \vee \overline{A}(x)),$$

from which the other follows by substituting $(Ez)A(x, z)$ for $A(x)$. But it is not recursively fulfillable, since it can be fulfilled only by the representing function of the predicate $(Ez)T_1(x, x, z)$, which, as we saw in the proof of Theorem II, is non-recursive.

17. **Non-elementary predicates.** The elementary predicates are enumerable. By Cantor's methods, there are therefore non-elementary number-theoretic predicates. However let us ask what form of definition would suffice to give such a predicate. Under classical interpretations, the enumeration of predicate forms given in Theorem II for n variables suffices for the expression of every elementary predicate of n variables. By defining relations of the form shown in the next theorem, we can introduce a predicate $M(a, k)$ so that it depends for different values of k on different numbers of alternating quantifiers. On the basis of Theorem II, it is possible to do this in such a way that the predicate will be expressible in none of the forms of Theorem II.

THEOREM XI. *Classically, there is a non-elementary predicate $M(a, k)$ definable by relations of the form*

$$\begin{cases} M(a, 0) \equiv R(a) \\ M(a, 2k+1) \equiv (Ex)M(\phi(a, x), 2k) \\ M(a, 2k+2) \equiv (x)M(\phi(a, x), 2k+1) \end{cases}$$

where R and ϕ are primitive recursive.

We are dealing here with essentially the same fact which Hilbert-Bernays discover by setting up a truth definition for their formal system (Z)(41).

The system (Z) has as primitive terms only $'$, $+$, $\cdot$, $=$ and the logical operations. The predicates expressible in these terms are elementary. Con-

(40) Heyting [1, p. 65].
(41) Hilbert and Bernays [1, pp. 328–340].

versely, using Theorem IV and Gödel's reduction of primitive recursive functions to these terms[42], every elementary predicate is expressible in (Z).

The Hilbert-Bernays result is an application to (Z) of Tarski's theorem on the truth concept[43], with the determination of a particular form of relations which give the truth definition for (Z). If (Z) is consistent, a formal proof that the relations do define a predicate is beyond the resources of (Z).

Bibliography

ALONZO CHURCH

1. *An unsolvable problem of elementary number theory*, Amer. J. Math. vol. 58 (1936) pp. 345–363.
2. *The constructive second number class*, Bull. Amer. Math. Soc. vol. 44 (1938) pp. 224–232.

ALONZO CHURCH AND S. C. KLEENE

1. *Formal definitions in the theory of ordinal numbers*, Fund. Math. vol. 28 (1936) pp. 11–21.

ALONZO CHURCH AND BARKLEY ROSSER

1. *Some properties of conversion*, Trans. Amer. Math. Soc. vol. 39 (1936) pp. 472–482.

H. B. CURRY

1. *Some aspects of the problem of mathematical rigor*, Bull. Amer. Math. Soc. vol. 47 (1941) pp. 221–241.
2. *The inconsistency of certain formal logics*, J. Symbolic Logic vol. 7 (1942) pp. 115–117.

KURT GÖDEL

1. *Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme* I, Monatshefte für Mathematik und Physik vol. 38 (1931) pp. 173–198.
2. *On undecidable propositions of formal mathematical systems*, notes of lectures at the Institute for Advanced Study, 1934.

DAVID HILBERT AND PAUL BERNAYS

1. *Grundlagen der Mathematik*, vol. 2, Berlin, Springer, 1939.

AREND HEYTING

1. *Die formalen Regeln der intuitionistischen Mathematik*, Preuss. Akad. Wiss. Sitzungsber, Phys.-math. Kl. 1930, pp. 57–71, 158–169.

S. C. KLEENE

1. *A theory of positive integers in formal logic*, Amer. J. Math. vol. 57 (1935) pp. 153–173, 219–244.
2. *General recursive functions of natural numbers*, Math. Ann. vol. 112 (1936) pp. 727–742.
3. *A note on recursive functions*, Bull. Amer. Math. Soc. vol. 42 (1936) pp. 544–546.
4. *On notation for ordinal numbers*, J. Symbolic Logic vol. 3 (1938) pp. 150–155.
5. *On the forms of the predicates in the theory of constructive ordinals*, to appear in Amer. J. Math. (Bull. Amer. Math. Soc. abstract 48-5-215).
6. *On the interpretation of intuitionistic number theory*, Bull. Amer. Math. Soc. abstract 48-1-85.

S. C. KLEENE AND BARKLEY ROSSER

1. *The inconsistency of certain formal logics*, Ann. of Math. (2) vol. 36 (1935) pp. 630–636.

DAVID NELSON

1. *Recursive functions and intuitionistic number theory*, under preparation.

E. L. POST

1. *Finite combinatory processes—formulation* I, J. Symbolic Logic vol. 1 (1936) pp. 103–105.

---

(42) Gödel [1, Theorem VII]. See Kleene [3 (erratum: p. 544, line 11, "of" should be at the end of the line)].

(43) Tarski [1].

W. V. QUINE

1. *Mathematical logic*, New York, Norton, 1940.

BARKLEY ROSSER

1. *Extensions of some theorems of Gödel and Church*, J. Symbolic Logic vol. 1 (1936) pp. 87–91.
2. *Gödel theorems for non-constructive logics*, ibid. vol. 2 (1937) pp. 129–137.

ALFRED TARSKI

1. *Der Wahrheitsbegriff in den formalisierten Sprachen*, Studia Philosophica vol. 1 (1936) pp. 261–405. (Original in Polish, 1933.)
2. *On undecidable statements in enlarged systems of logic and the concept of truth*, J. Symbolic Logic vol. 4 (1939) pp. 105–112.

A. M. TURING

1. *On computable numbers, with an application to the Entscheidungsproblem*, Proc. London Math. Soc. (2) vol. 42 (1937) pp. 230–265.
2. *Systems of logic based on ordinals*, ibid. vol. 45 (1939) pp. 161–228.

AMHERST COLLEGE,
AMHERST, MASS.

BÉZOUT'S THEOREM AND ALGEBRAIC DIFFERENTIAL EQUATIONS(1)

BY

J. F. RITT

The problem of determining by inspection the number of solutions of a system of algebraic equations finds its solution in Bézout's theorem and in important complements to that theorem obtained in recent years by van der Waerden(2). The corresponding problem for a system of algebraic differential equations is that of determining bounds for the numbers of arbitrary constants which enter into the irreducible manifolds which the system yields. This problem has been considered by us in two previous papers(3).

In the present paper, we study the intersections of the ***general solutions*** of two algebraically irreducible forms A and B in the unknowns y and z. The statement of our results depends on some definitions which we proceed to give.

Let F be a form in several unknowns. F has an order in each of its unknowns. The maximum of these orders will be called the ***order of F***.

Let Σ be a non-trivial prime ideal of forms in any unknowns. Σ has a certain number $q \geqq 0$ of arbitrary unknowns. We shall call q the ***dimension*** of the manifold of Σ.

By the order of an irreducible manifold $\mathfrak{M}$ of dimension zero, we mean the order of any resolvent for the prime ideal of which $\mathfrak{M}$ is the manifold.

An irreducible manifold $\mathfrak{N}$ which is part of a manifold $\mathfrak{M}$ will be called an ***irreducible component*** (often simply component) of $\mathfrak{M}$ if $\mathfrak{M}$ contains no irreducible manifold of which $\mathfrak{N}$ is a proper part(4).

Let us return now to A and B as above, which we suppose to have the respective orders m and n. Let the general solutions of A and B have a non-vacuous intersection $\mathfrak{M}$. It is a most natural conjecture that, if $\mathfrak{M}$ has one or more irreducible components of dimension zero, their orders do not exceed $m+n$. This conjecture is verified below for the cases in which neither of m and n exceeds unity. It was not without surprise that we found our conjecture to lapse into default for larger values of the orders. We shall show how to construct, for every $n \geqq 4$, a form of order n whose general solution intersects

---

Presented to the Society, October 31, 1942; received by the editors March 23, 1942.

(1) For indications in regard to the general theory to which this paper attaches, one may consult the author's paper in the second volume of the Semicentennial Publications of the American Mathematical Society.

(2) van der Waerden, *Einführung in die algebraische Geometrie*, Berlin, 1939, chap. 6.

(3) *Systems of algebraic differential equations*, Ann. of Math. (2) vol. 36 (1935) p. 293; *Jacobi's problem on the order of a system of differential equations*, ibid. p. 303. The second of these papers will be denoted below by J.

(4) In other words, $\mathfrak{N}$ is essential in $\mathfrak{M}$.

the manifold of $y=0$ in the manifold of $y=0$, $z_{2n-3}=0$, a manifold of order $2n-3$.

FORMS OF ORDERS NOT EXCEEDING UNITY

1. We prove the statement made, for the cases with $m\leqq 1$, $n\leqq 1$, in the introduction.

When $m=n=0$, there is nothing to prove.

Let $m=0$, $n=1$. Let $\mathfrak{N}$ be a component of $\mathfrak{M}$ of dimension zero. We consider first the intersection $\mathfrak{M}'$ of the complete manifolds of A and B. Every component of $\mathfrak{M}'$ of dimension zero has an order not exceeding unity(5). Then, by Gourin's theorem(6), if $\mathfrak{N}$ is not contained in a component of $\mathfrak{M}'$ of dimension unity, $\mathfrak{N}$ is of order not greater than unity.

We have now to consider the case in which $\mathfrak{N}$ is contained in a component $\mathfrak{M}''$ of $\mathfrak{M}'$ of dimension unity. $\mathfrak{M}''$ is the general solution of a form C. Because A, which is of order zero, holds $\mathfrak{M}''$, C must be of order zero; this implies that $\mathfrak{M}''$ is the manifold of A. Then $\mathfrak{M}''$ must be a component of the manifold of B. Otherwise $\mathfrak{M}''$ would be contained in the general solution of B and $\mathfrak{N}$ would not be a component of $\mathfrak{M}$(7).

We suppose, as we may, that A involves z effectively. As $\mathfrak{M}''$ is a proper part of the manifold of B, B must be of order unity in z. Let $S=\partial A/\partial z$. Then some S^tB has a representation(8)

$$S^tB = C_0A^p + C_1A^{p_1}A_1^{q_1} + \cdots + C_rA^{p_r}A_1^{q_r}.$$

Here A_1 is the derivative of A and, for every i, $p_i+q_i>p$. The orders of the C_i in z and in y do not exceed 0 and 1, respectively, and no C_i is divisible by A. As $\mathfrak{N}$ is in the intersection of $\mathfrak{M}''$ and the general solution of B, C_0 must hold $\mathfrak{N}$(8). The manifold of the system C_0, A has components which are all of dimension zero and none of order greater than unity(8). This disposes of the case of $m=0$, $n=1$.

Now let $m=n=1$. We use $\mathfrak{N}$ and $\mathfrak{M}'$ as above. We take up immediately the case in which $\mathfrak{N}$ is contained in a component $\mathfrak{M}''$ of $\mathfrak{M}'$ of dimension unity; when $\mathfrak{N}$ is not so contained, its order cannot exceed 2(9). As $\mathfrak{N}$ is a component of $\mathfrak{M}$, $\mathfrak{M}''$ is not part of $\mathfrak{M}$. Let, then, $\mathfrak{M}''$ fail to be contained in the general solution of B. Then some other component of the manifold of B, indeed the manifold of a form of order zero, contains $\mathfrak{M}''$ and is thus identical with $\mathfrak{M}''$. By the case of $m=0$, $n=1$, the components of the inter-

(5) This is proved in J.

(6) Bull. Amer. Math. Soc. vol. 39 (1933) p. 593.

(7) The components of B other than its general solution are manifolds of forms of order zero. See *On certain points in the theory of algebraic differential equations*, Amer. J. Math. vol. 60 (1938) pp. 1–43, §30. This paper will be denoted by C. P.

(8) C. P. §31.

(9) By J.

section of $\mathfrak{M}''$ with the general solution of B are of dimension zero and of order at most unity. This completes the proof.

A FORM OF ORDER FOUR

2. In what follows, K_1 will represent, for any form K, the derivative of K. We let

$$A = y_1 - z_3 y^2, \tag{1}$$

$$B = A^4 - y_3^8, \tag{2}$$

$$C = y_3 A_1 - 2y_4 A, \tag{3}$$

$$F = B - y^6 C^2 = A^4 - y_3^8 - y^6 C^2. \tag{4}$$

We use the field of all constants. Let us see first that F is algebraically irreducible. If we consider the equation $F=0$ as an algebraic equation for y_4, we secure a function y_4 of two branches. Thus, if F were factorable, it would have a factor of positive degree free of y_4. Such a factor would have to be a factor of $y^6 A^2$. As F is not divisible by y or by A, F is algebraically irreducible.

Let us determine now the components of the manifold of F other than the general solution.

Let $\mathfrak{N}$ be such a component. As $\partial F/\partial y_4 = 4y^6 AC$, $\mathfrak{N}$ must be held by yC or by A. Suppose that A holds $\mathfrak{N}$. By (3) and (4), y_3 holds $\mathfrak{N}$. In every case then, B holds $\mathfrak{N}$.

Now B is the product of the four forms

$$E^{(j)} = y_1 - z_3 y^2 - j y_3^2, \qquad j = \pm 1, \pm (-1)^{1/2}, \tag{5}$$

each of which is algebraically irreducible. For what follows, it is important to know that the manifold of each $E^{(j)}$ is irreducible. From the manner in which z_3 figures in (5), one sees that a component of the manifold of an $E^{(j)}$ distinct from the general solution is held by y. Such a component, being of dimension unity(10), must be the manifold of y. But the ***low power theorem***(11) shows that the manifold of y is not a component. This proves the irreducibility of the manifolds of the $E^{(j)}$.

We have, for every j,

$$C = y_3 E_1^{(j)} - 2y_4 E^{(j)}.$$

Referring to (4) and applying the low power theorem, we find that the manifold of each $E^{(j)}$ is a component of the manifold of F(12).

---

(10) C. P. §1.

(11) So we designate the theorem of C. P. §29.

(12) Technically, in applying the low power theorem, we have to multiply F by y_3^2 and to effect a reduction. Actually, on considering the proof of the low power theorem, one sees that one may dispense with this process of preparation. For instance, if one replaces y_4 in the coeffi-

Thus the manifold of F has five components, the general solution and the manifolds of the $E^{(j)}$.

3. In what follows it will be proved that the intersection of the general solution of F with the manifold of $y=0$, is the manifold of the system $y=0$, $z_5=0$. The latter manifold is of dimension zero and of order 5. The proof employs some general results, bearing on ideals of differential polynomials, which will now be set forth.

Deductions from Levi's theorem on power products

4. In what follows P is a power product in y and derivatives of y, d the degree of P, w the weight of P and p a positive integer.

Modifying a theorem due to Howard Levi([13]), we derive the following result: *If*

$$d > \frac{p-1}{2} + \left((p-1)w + \frac{(p-1)^2}{4}\right)^{1/2} \tag{6}$$

then([14])

$$P \equiv 0, \qquad [y^p].$$

We suppose, as we may, that $p>1$. Let (6) be satisfied. Then

$$(p-1)w < d^2 - d(p-1). \tag{7}$$

Let

$$d = a(p-1) + b \tag{8}$$

where a and b are integers such that $a \geqq 0$, $0<b\leqq p-1$. As $b(p-1-b)\geqq 0$, (7) gives

$$(p-1)w < d^2 - d(p-1) + (p-1-b)b. \tag{9}$$

We replace d in (9) by its expression in (8), finding that

$$w < a(a-1)(p-1) + 2ab. \tag{10}$$

By Levi's theorem, $P\equiv 0$, $[y^p]$.

We denote by $\delta(p, w)$ the second member of (6).

5. Representing y^p by u, we prove the following result, which holds for any power product P as in §4 and for any values of d, w, p.

P has a representation as a homogeneous polynomial in u and derivatives of u,

---

cient of $E^{(j)}$ in C by a new unknown u, $E^{(j)}$ is seen immediately to furnish a component of the manifold of the form in u, y, z into which F is converted.

([13]) Trans. Amer. Math. Soc. vol. 51 (1942) p. 545.

([14]) The notation, as regards congruences, is due to E. R. Kolchin, Ann. of Math. (2) vol. 42 (1941) p. 740.

whose coefficients are homogeneous polynomials[15] *in y and derivatives of y of a common degree not greater than $\delta(p, w)$.*

If $d \leqq \delta(p, w)$, P itself is the representation sought. Otherwise, by §4, P is a linear combination of the u_i, with coefficients all of degree $d-p$ and none of weight greater than w. If $d-p \leqq \delta(p, w)$, we have the desired representation. Otherwise the coefficients of the u_i will be in $[u]$. Continuing in this manner, we find P expressed as in our statement.

Multipliers of a form

6. Let Σ be an ideal (differential) of forms in y and z; M a form in y and z; α a non-negative number. We shall say that *M admits α as a multiplier with respect to Σ* if for every $\epsilon > 0$ there exists an integer $n_0(\epsilon)$ such that, for every $n > n_0(\epsilon)$,

$$M^n \equiv P, \qquad [\Sigma]$$

where P is a form depending on n which, arranged as a polynomial in the y_i[16], contains no term of degree less than $n(\alpha-\epsilon)$. P may be zero. If α is a multiplier for M and if $0 \leqq \gamma < \alpha$, γ is also a multiplier.

We prove the following properties of multipliers:

(a) Let M and N admit α and β, respectively, as multipliers with respect to Σ. Let $\gamma = \min(\alpha, \beta)$. Then $M+N$ admits γ as a multiplier.

(b) For M and N as in (a), MN admits $\alpha+\beta$ as a multiplier.

(c) Let M^p, where p is a positive integer, admit α as a multiplier. Then M admits α/p.

(d) Let M admit α as a multiplier. Then M_1, the derivative of M, also admits α.

(e) If $M \equiv N$, $[\Sigma]$, M and N admit the same multipliers.

Proving (a), we take an $\epsilon > 0$. Let $n_0(\epsilon/2)$ serve as above for both M and N with respect to $\epsilon/2$. We consider $(M+N)^n$ for any $n \geqq 1$. Let $R = M^a N^b$ where $a+b=n$. If a and b both exceed $n_0(\epsilon/2)$, we have $R \equiv P$, $[\Sigma]$ where no term of P is of degree less than

$$a(\alpha - \epsilon/2) + b(\beta - \epsilon/2),$$

which quantity is not less than $n(\gamma - \epsilon/2)$. If $b \leqq n_0(\epsilon/2) < a$, we have $R \equiv P$, $[\Sigma]$ with no term of P of degree less than

$$[n - n_0(\epsilon/2)](\alpha - \epsilon/2).$$

This last quantity, if n is large in comparison with $n_0(\epsilon/2)$, exceeds $n(\alpha-\epsilon)$. The truth of (a) is now clear.

---

(15) The coefficients of the polynomials in the y_i are rational numbers.

(16) When P is thus arranged, its coefficients will be forms in z. The definition of multiplier gives a special role to y.

The proofs of (b), (c) and (e) are trivial.

Proving (d), we take an $\epsilon>0$ and, relative to M, an $n_0(\epsilon/2)$. Let m be a fixed integer which exceeds $n_0(\epsilon/2)$. We consider an $n>0$ and use $\delta(m, n)$ as in §4. Then M_1^n is a polynomial in M^m and its derivatives with coefficients which are forms in M of degree not greater than $\delta(m, n)$. In this expression for M_1^n, every power product in M^m and its derivatives is of degree not less than

$$q = [n - \delta(m, n)]/m. \tag{11}$$

Now, if n is large, $\delta(m, n)$ as one sees from (6), is small in comparison with n, so that q is only slightly less than n/m. Each power product in M^m and its derivatives is congruent to a form whose terms have degrees in the y_i not less than $qm(\alpha-\epsilon/2)$. If n is large, this last quantity exceeds $n(\alpha-\epsilon)$, q.e.d.

The form F. First operation

7. We return to F of §2, denoting the general solution of F by $\mathfrak{M}$. We show now that a solution in $\mathfrak{M}$ with $y=0$ satisfies $z_5=0$. Later, we shall prove that every z with $z_5=0$ is admissible.

We determine first a form G which holds $\mathfrak{M}$ but none of the other four components.

We have by (2) and (3),

$$AB_1 - 4A_1B = 4y_3^7C. \tag{12}$$

Thus by (4) (first representation of F), we have, when $F=0$,

$$4y_3^7B^{1/2} = y^3(AB_1 - 4A_1B). \tag{13}$$

Again, letting $K=y^3C$, we have by (4), when $F=0$, the relation $B^{1/2}=K$. Thus, for $F=0$, $B\neq 0$,

$$B^{-1/2}B_1 = 2K_1. \tag{14}$$

Substituting into (13) the expression which (14) furnishes for B_1, and simplifying, we find for $F=0$, $B\neq 0$,

$$4y_3^{14} + L = 0 \tag{15}$$

where

$$L = - 4y^3y_3^7AK_1 + y^6A^2K_1^2 - 4y^6A_1^2B. \tag{16}$$

We designate the first member of (15) by G. Then G holds $\mathfrak{M}$.

8. In what follows, all multipliers will operate with respect to $[F, G]$, the differential ideal generated by F and G.

In (4), y_3^8 and y^6C^2 contain no terms of degree less than 8 in the y_i. Thus A^4 admits 8 as a multiplier so that, by (c) of §6, A admits 2. Now z_5y^2 admits 2. By (a) of §6, y_1 admits 2. Then, by (d), every y_i with $i\geq 1$ admits 2. From

(3), using (a), (b), (d), we find that C admits 4. Referring to (4) and using (e), we see now that A^4 admits 14, so that A admits 3. By (3), now, C admits 5 and we find from (4) that A admits 4. We return to (3) and see that C admits 6. Also by (4), B admits 18. Finally K of §7 admits 9.

By (16), L admits 30. By (15), y_3 admits 15/7. Now $y_2-z_4y^2-2z_3yy_1$, which is A_1, admits 4. As y_1 admits 2, $y_2-z_4y^2$ admits 3. Then $y_3-z_5y^2-2z_4yy_1$ admits 3, so that $y_3-z_5y^2$ admits 3. As y_3 admits 15/7, z_5y^2 admits 15/7.

We infer that $[F, G]$ contains a form of the type $(z_5y^2)^m+M$ where every term of M is of degree greater than $2m$ in the y_i. It follows from the low power theorem that a solution in $\mathfrak{M}$ cannot have $y=0$ unless $z_5=0$.

Second operation

9. Let α be any polynomial of effective degree 4. We shall prove that $\mathfrak{M}$ contains $y=0$, $z=\alpha$. This will imply that every z for which $z_5=0$ appears in $\mathfrak{M}$ with $y=0$ and our investigation of F will be completed.

Representing by c an arbitrary constant and by v a new unknown, we put $z=\alpha$ in F and then make in F the substitution[17]

$$y = \sum_{j=1}^{6} c^{j}\alpha_2^{j-1} + c^6 v. \tag{17}$$

We represent by A', A_1', B', C', F' the expressions into which A, A_1, B, C, F are, respectively, transformed when z is replaced by α and y by the second member of (17).

We find from (17)

$$A' = c^6v_1 + c^7P \tag{18}$$

with P a polynomial in x, c, v. Then we may write

$$A_1' = c^6v_2 + c^7Q. \tag{19}$$

In (17), the coefficient of c^2 is of the second degree in x; that of c^3 is of the fourth degree. We have thus

$$y_3 = c^3\beta + \cdots; \qquad y_4 = c^3\gamma + \cdots \tag{20}$$

with β of the first degree and γ constant. By (18), (19), (20),

$$C' = c^9(\beta v_2 - 2\gamma v_1) + c^{10}R$$

with R a polynomial in x, c and the v_j with $j\leqq 4$. We find thus

$$F' = c^{24}\left[v_1^4 - \beta^8 - (\beta v_2 - 2\gamma v_1)^2\right] + c^{25}T \tag{21}$$

with T of the type of R.

---

[17] Subscripts of α indicate differentiation.

10. Let V represent the coefficient of c^{24} in F'. As $\beta \neq 0$, the differential equation $V=0$ is effectively of the second order. Let then $v=\xi$ be a solution of $V=0$ with

$$\xi_1^4 - \beta^8 \neq 0. \tag{22}$$

We wish to show that F' is formally annulled by a series

$$v = \xi + \phi_2 c^{\rho_2} + \phi_3 c^{\rho_3} + \cdots \tag{23}$$

of the following description. The ρ_i are positive rational numbers, with a common denominator, which increase with their subscripts. The ϕ_i are analytic functions of x, all analytic at some point at which ξ is analytic[18].

It will suffice to show that $G=F'/c^{24}$ is annulled by a series (23). If G vanishes identically in x and c for $v=\xi$, then $v=\xi$ is an acceptable series (23). In what follows, we assume that such vanishing does not occur.

Introducing a new unknown u_1, we put, in G, $v=\xi+u_1$. Then G goes over into an expression H' in x, c and u_1,

$$H' = a'(c) + \sum b_i'(c) u_{10}^{\alpha_{0i}} \cdots u_{14}^{\alpha_{4i}}. \tag{24}$$

Here $\sum$ contains the terms of H' which are not free of the u_{1j} and, in $\sum$, i ranges from unity to some positive integer. As to $a'(c)$ and the $b_i'(c)$, they are polynomials in c with analytic functions of x for coefficients. Because ξ does not annul G identically, $a'(c)$ is not identically zero. On the other hand, because G vanishes for $v=\xi$, $c=0$, the lowest power of c in $a'(c)$ is positive. Because the bracketed terms in (21) contribute effectively to $\sum$ in (24), certain of the $b_i'(c)$ contain terms of power zero in c.

Let σ' be the least exponent of c in a' and σ_i' the least exponent of c in b_i'. Let

$$\rho_2 = \max \frac{\sigma' - \sigma_i'}{\alpha_{0i} + \cdots + \alpha_{4i}}$$

where i has the range which it has in $\sum$. As $\sigma'>0$ and certain σ_i' equal 0, $\rho_2>0$.

We now take over §§12–16 of our paper *On the singular solutions of algebraic differential equations*[19], putting $m=4$ in that discussion. We are brought to the series (23) for v.

11. We have shown, all in all, that F, for $z=\alpha$, is annulled by a series

$$y = c + c^2\alpha_2 + \cdots + c^5\alpha_2^4 + c^6(\alpha_2^5 + \xi) + \cdots \tag{25}$$

where the unwritten terms have rational exponents greater than 6. The series (25) does not annul B for $z=\alpha$. Indeed,

---

(18) One may suppose that $\phi_1=\xi$, $\rho_1=0$.

(19) Ann. of Math. (2) vol. 37 (1936) p. 541.

$$B' = c^{24}(v_1^4 - \beta^8) + \cdots$$

and, because of (22), the coefficient of c^{24} in B' does not vanish for $v=\xi$.

It follows that every form which holds $\mathfrak{M}$ vanishes for $z=\alpha$ and for y as in (25). This means that $y=0$, $z=\alpha$ is in $\mathfrak{M}$.

Remarks

12. If in (1) to (4), we replace z_3, y_3, y_4 wherever they appear by z_{n-1}, y_{n-1}, y_n, respectively, where $n \geqq 4$, we obtain a form F with a general solution which intersects the manifold of $y=0$ in that of $y=0$, $z_{2n-3}=0$; the proofs require only the slightest changes.

In F of §2, if one replaces z_3 by z, one obtains a form which is of the first order in z and has a general solution which intersects the manifold of $y=0$ in that of $y=0$, $z_3=0$. This in itself is sufficiently anomalous. However, if it is desired to secure a form F whose order in z cannot be reduced, it suffices to replace y_3 and y_4 in (2), (3), (4) by zy_3 and its derivative, respectively.

Columbia University,
New York, N. Y.

THE CONTINUITY OF FUNCTIONS OF MANY VARIABLES

BY

RICHARD KERSHNER

1. **Introduction.** It is known that a function $f(x, y)$ of two real variables may be continuous with respect to each variable separately throughout a given region without being continuous with respect to (x, y) at all points of the region. In fact, W. H. and G. C. Young(1) have given an example of a function $f(x, y)$ which is a continuous function of the position along every straight line in the unit square $[0, 1]\times[0, 1]$ but which has an uncountable number of two-dimensional discontinuities in every rectangle contained in the unit square. The example of W. H. and G. C. Young could easily be modified so as to yield a function continuous along every analytic arc but with an uncountable number of discontinuities in every rectangle.

If the number of variables is greater than two the situation becomes even worse. As was pointed out by Baire(2), for three variables, and subsequently by Hahn(3), for any number of variables, a function $f(x_1, x_2, \cdots, x_n)$ may be continuous in each variable x_i and yet be discontinuous with respect to $(x_1, x_2, \cdots, x_n)$ at every point of an $(n-2)$-dimensional rectangle. In fact let $g(x_1, x_2)$ be a function continuous in x_1 and x_2 separately but discontinuous at $(0, 0)$. Then

$$f(x_1, x_2, \cdots, x_n) \equiv g(x_1, x_2)$$

is discontinuous at every point of the $(n-2)$-dimensional region $x_1=0$, $x_2=0$. Finally, Lebesgue(4) has shown that a function $f(x_1, x_2, \cdots, x_n)$ which is continuous in each variable x_i separately may be of the $(n-1)$st Baire class, although no worse.

The problem of considering how much could be said concerning the n-dimensional continuity points of a function $f(x_1, x_2, \cdots, x_n)$ which is assumed to be continuous in each x_i separately was first treated in 1899 by Baire in the fundamental paper(5) in which he introduced most of the classic notions associated with his name. For the case of two variables his results were complete. He showed that

---

Presented to the Society, December 30, 1941; received by the editors April 16, 1942.

(1) W. H. Young and G. C. Young, *Discontinuous functions continuous with respect to every straight line*, Quart. J. Math. Oxford Ser. vol. 41 (1910) pp. 87–93.

(2) R. Baire, *Sur les fonctions de variables réelles*, Annali di Mathematica Pura ed Applicata vol. 3 (1899) pp. 1–122.

(3) Hahn, *Über Funktionen mehrerer Veränderlichen, die nach jeder einzelnen Veränderlichen stetig sind*, Math. Zeit. vol. 4 (1919) pp. 306–313.

(4) H. Lebesgue, *Sur les fonctions représentable analytiquement*, J. Math. Pure Appl. (6) vol. 1 (1905) pp. 139–212, cf. pp. 201, 202.

(5) Loc. cit. Footnote 2.

(A) *If $f(x, y)$ is, in the unit square, a continuous function of x, for fixed y, and a continuous function of y, for fixed x, then there is a residual set of lines parallel to each axis consisting entirely of points where $f(x, y)$ is continuous with respect to (x, y).*

Here and throughout the paper a set contained in an interval is called a residual set if its complement in that interval is of first category.

It should be recalled that, in consequence of Baire's classic theorem, a residual set in an interval is uncountably dense in every subinterval. Thus the statement obtained from (A) by reading "dense" for "residual" is an immediate consequence of (A). Actually Baire explicitly displayed only this weaker consequence although he actually proved (A). It follows immediately from (A) that

(B) *Under the assumptions of* (A) *every line parallel to an axis contains a dense set of points where $f(x, y)$ is continuous with respect to (x, y).*

Clearly (A) is much stronger than (B). For example, (A) does, but (B) does not, imply that

(C) *Under the assumptions of* (A) *the set of points where $f(x, y)$ is discontinuous with respect to (x, y) has dimension (Menger) at most zero.*

Of course this beautiful theorem (C) was not stated by Baire, since the general notion of dimension involved is of a later date.

Baire also treated the case of three variables in the same 1899 paper. Unfortunately he was not able to generalize the result (A) (which, as is shown in §6, is definitive) but only the weaker consequence (B). Of course, in view of the fact that $f(x, y, z)$ may be of the *second* Baire class, it is rather surprising that even (B) may be extended to the case of three variables. However, Baire was able to show that every plane parallel to a coordinate plane must contain a dense set of continuity points for $f(x, y, z)$, assumed continuous in x, y and z separately.

The methods of Baire apparently yielded nothing for $f(x_1, x_2, \cdots, x_n)$ if $n>3$, and it was not until 1910 that Hahn(6) showed that, in spite of the increasing Baire class, a generalization of (B) remained valid for any n. In fact, it was shown that any $(n-1)$-dimensional hyperplane obtained by fixing one coordinate must contain a dense set of continuity points of a function $f(x_1, x_2, \cdots, x_n)$ assumed continuous in each x_i.

The only other consideration of the problem of which I am aware is a paper of Bogel(7) who established the conclusion of (B) under somewhat weaker assumptions than those of Baire.

---

(6) Loc. cit. Footnote 3.

(7) K. Bogel, *Über die Stetigkeit und die Schwankung von Funktionen zweier reeler Veranderlichen*, Math. Ann. vol. 81 (1920) pp. 64–93.

The main purpose of this paper is to show that not only (B) but also the definitive result (A) can be generalized to the case of an arbitrary number of variables. The precise result, stated here for the case of three variables, is the following:

If $f(x, y, z)$ is, in the unit cube, a continuous function of x, and of y, and of z then there is a residual set of planes parallel to each coordinate plane, on each of which there is a residual set of lines parallel to each possible coordinate axis consisting entirely of points where $f(x, y, z)$ is continuous with respect to (x, y, z).

In particular the result to be established in the case of an arbitrary number of variables is strong enough to establish the natural generalization of (C), that is,

If $f(x_1, x_2, \cdots, x_n)$ is, in the unit cube, a continuous function of each x_i, then the set of points where $f(x_1, x_2, \cdots, x_n)$ is discontinuous with respect to $(x_1, x_2, \cdots, x_n)$ has dimension at most $(n-2)$.

It will be recalled that the set of discontinuities may have dimension $n-2$ even if $f(x_1, x_2, \cdots, x_n)$ is an algebraic function.

The extension of the Baire results to the case of more than two variables is based mainly on a set-theoretic lemma (Lemma 2) which is proved in §2. Section 2 also contains a list of the notations and results of a set-theoretic nature that will be needed later.

Section 3 contains a similar list of notations and results of a function-theoretic nature that will be used. These are surprisingly few in number and elementary in nature.

In §4 a greatly simplified proof is given of the Baire result for a function of two variables. This is technically unnecessary since the induction proof to be given in §5 for the case of any number of variables could be based on the completely trivial case $n=1$ rather than the Baire case $n=2$, but the Baire result does not seem to be so well known as it deserves to be and it seemed that a direct modern proof might be useful.

In §6 it is shown that the results of §5 are the best possible. It seems not to have been known even that the Baire result (A) was best possible.

It might be wondered whether the results of §5 can be strengthened by requiring more smoothness, for example, the existence of partial derivatives or a 1-dimensional Lipschitz condition, parallel to the axes. This problem is treated in §7, where it is shown that an assumption considerably weaker than a 1-dimensional Lipschitz condition of any order $\alpha>0$ is sufficient to ensure that the discontinuities of $f(x_1, x_2, \cdots, x_n)$ are nowhere dense. This result seems to be new even in the simplest case $n=2$, although Bogel([8]) has a much weaker result in this direction in the case $n=2$.

Finally, in §8, it is shown that the result of §7 is definitive and in fact that no further restrictions on the set of discontinuities are imposed by requiring any degree of smoothness, short of analyticity, parallel to the axes.

([8]) K. Bogel, *Über partiell differenzierbare Funktionen*, Math. Zeit. vol. 25 (1926) pp. 490–498.

It might be mentioned that the results of this paper could easily be extended to quite abstract product spaces but the author's interests do not lie in that direction.

2. **Preliminaries.** In this section will be listed certain notations and results, some of them classic, that will be used in this paper. Special attention is called to the set-theoretic Lemma 2, which seems to be rather powerful.

Set. The bracket notation $[a; \cdots]$ (or $[a \in S; \cdots]$), is used to denote the set of all those elements a (of S) for which the specified restrictions "$\cdots$" are satisfied.

Interval. All intervals are understood to be open unless otherwise specified. The closed unit real interval $[0, 1]$ is denoted by U or U_1. The closed unit n-interval $U \times U \times \cdots \times U$ is denoted by U_n. If $I = (a, b)$ is a subinterval of U then $|I|$ denotes its length $b - a$.

F_σ. A set is called an F_σ-set if it is the sum of a countable number of closed sets.

Dimension. The recursive (Menger-Urysohn) definition of dimension will be used. The empty set has dimension -1. A set S has dimension at most m if each neighborhood of every point of S contains another neighborhood of that point whose boundary intersects S in a set with dimension at most $m-1$.

Category. A subset of U_n is called of the first category if it is the sum of a countable number of sets nowhere dense in U_n. All other subsets of U_n are called of the second category.

Residual. A subset of an interval of U_n is called residual in that interval if its complement in that interval is of the first category.

BAIRE'S THEOREM. *The empty set is not residual, that is, an interval is of the second category.*

Baire's theorem will be used mainly in conjunction with one of the following two lemmas of which the first is classic (and trivial) and the second seems to be new.

LEMMA 1. *Let $\{S_j\}$ be a sequence of sets in U_n such that S_j is closed and $\sum S_j$ is of the second category in U_n. Then, for some integer k, S_k contains a subinterval of U_n.*

Proof. Since $\sum S_j$ is of the second category, not all S_j are nowhere dense. Thus some S_k is dense in a subinterval I of U_n. But S_k is closed, so $S_k \supset I$.

In order to facilitate the statement of the other desired lemma, two definitions are given next.

Linearly closed. A set $S \subset U_n$ will be called linearly closed if it intersects each line parallel to a coordinate axis of U_n in a closed set. Thus S is linearly closed if, for each fixed

$$(x_1, x_2, \cdots, x_{i-1}, x_{i+1}, \cdots, x_n) \in U_{n-1}$$

the set

$$[x_i \in U; (x_1, x_2; \cdots, x_{i-1}, x_i, x_{i+1}, \cdots, x_n) \in S]$$

is closed.

Clearly a closed set is linearly closed, but not conversely.

Property Δ. This will be defined for sets in U_n inductively with respect to n. A set $S \subset U_1$ has property Δ if it is of second category in U_1. A set $S \subset U_n$ has property Δ if there is a set $R \subset U_{n-1}$ such that

$$R \text{ has property } \Delta \text{ in } U_{n-1}$$

and such that if

$$\mathbf{x}_{n-1} \equiv (x_1, x_2, \cdots, x_{n-1}) \in R$$

then the set

$$[x_n \in U; (\mathbf{x}_{n-1}, x_n) \in S]$$

is of second category in U.

In view of Baire's theorem it is clear that U_n itself has property Δ and R may be chosen as U_{n-1}. More generally any set which contains a subinterval of U_n has property Δ in U_n.

Lemma 2. *Let* $\{S_j\}$ *be a sequence of sets in* U_n *such that* S_j *is linearly closed and* $\sum S_j$ *has property* Δ. *Then, for some* k, S_k *contains a subinterval of* U_n.

Proof. The proof is by induction on n. If $n=1$, Lemma 2 follows from Lemma 1. Suppose then that the case $n-1$ of Lemma 2 is valid.

Since $\sum S_j$ has property Δ there is a set $R \subset U_{n-1}$ such that

$$R \text{ has property } \Delta \text{ in } U_{n-1} \tag{1}$$

and

$$\mathbf{x}_{n-1} \in R \text{ implies } T \equiv [x_n; (\mathbf{x}_{n-1}, x_n) \in \textstyle\sum S_j] \text{ of second category.} \tag{2}$$

For fixed $\mathbf{x}_{n-1} \in R$ let

$$T_j \equiv T_j(\mathbf{x}_{n-1}) \equiv [x_n; (\mathbf{x}_{n-1}, x_n) \in S_j], \qquad j = 1, 2, \cdots. \tag{3}$$

Then T_j is closed since S_j is linearly closed. Also $\sum T_j = T$ is of second category by (3), (2). Thus, by Lemma 1, there is, for each $\mathbf{x}_{n-1} \in R$, at least one integer k and a corresponding interval $I \subset U$ such that $T_k \supset I$. Thus, if for each $\mathbf{x}_{n-1} \in R$ and for each $k = 1, 2, \cdots$, one defines

$$\lambda(\mathbf{x}_{n-1}, k) \equiv \max [\,|I|\,; I \subset T_k(\mathbf{x}_{n-1}) \subset U] \tag{4}$$

(which exists since T_k is closed), then

$$\text{for every } \mathbf{x}_{n-1} \in R \text{ there is a } k \text{ such that } \lambda(\mathbf{x}_{n-1}, k) > 0. \tag{5}$$

Now, for each $j=1, 2, \cdots; k=1, 2, \cdots$, let

$$M_{j,k} \equiv [\mathbf{x}_{n-1} \in R; \lambda(\mathbf{x}_{n-1}, k) \geqq 1/j].$$

Then, clearly, $\mathbf{x}_{n-1} \in M_{j,k}$ if and only if there is an interval $I = I(\mathbf{x}_{n-1}, k) \subset U$, such that

$$(6) \qquad |I(\mathbf{x}_{n-1}, k)| \geqq 1/j \text{ and } x_n \in I(\mathbf{x}_{n-1}, k) \text{ implies } (\mathbf{x}_{n-1}, x_n) \in S_k.$$

It will now be shown that the induction assumption may be applied to the sequence of sets $M_{j,k} \subset R \subset U_{n-1}$. First $\sum M_{j,k} = R$. In fact if $\mathbf{x}_{n-1} \in R$ then by (5), $\lambda(\mathbf{x}_{n-1}, k) \geqq 1/j$ for some j, k, and so $\mathbf{x}_{n-1} \in \sum M_{j,k}$. Thus $\sum M_{j,k}$ has property Δ by (1).

It remains to show that $M_{j,k}$ is linearly closed. To this end it is sufficient to prove that if $\{x_{i,h}\}$ is, for fixed $1 \leqq i \leqq n-1$, a sequence of numbers in U such that

$$(7) \qquad \lim_{h \to \infty} x_{i,h} = \bar{x}_i \text{ exists}$$

and

$$(8) \qquad (x_1, x_2, \cdots, x_{i-1}, x_{i,h}, x_{i+1}, \cdots, x_{n-1}) \in M_{j,k}, \qquad h = 1, 2, \cdots,$$

then

$$(x_1, x_2, \cdots, x_{i-1}, \bar{x}_i, x_{i+1}, \cdots, x_{n-1}) \in M_{j,k}.$$

It has been seen, in (6), that (8) implies the existence of an interval $I = I_h \subset U$ such that $|I_h| \geqq 1/j$ and

$$x_n \in I_h \text{ implies } (x_1, x_2, \cdots, x_{i-1}, x_{i,h}, x_{i+1}, \cdots, x_{n-1}, x_n) \in S_k.$$

Let $x_{n,h}$ denote the midpoint of I_h. Then there is a subsequence x_{n,h_m} which is convergent, say

$$x_{n,h_m} \to \bar{x}_n.$$

Now let x_n be a fixed number in U such that

$$(9) \qquad |x_n - \bar{x}_n| < 1/2j$$

so that

$$|x_n - \bar{x}_n| = (1/2j) - \delta$$

for some $\delta > 0$. Then, for sufficiently large m,

$$|x_{n,h_m} - \bar{x}_n| < \delta$$

so that

$$|x_n - x_{n,h_m}| < 1/2j.$$

Thus $x_n \in I_{h_m}$ for sufficiently large m. Then

$$(x_1, x_2, \cdots, x_{i-1}, x_{i,h_m}, x_{i+1}, \cdots, x_{n-1}, x_n) \in S_k,$$

for sufficiently large m. Then, since S_k is linearly closed,

$$(x_1, x_2, \cdots, x_{i-1}, \bar{x}_i, x_{i+1}, \cdots, x_{n-1}, x_n) \in S_k$$

by (7). But this is true for any fixed x_n in the interval (9) of length $1/j$. Thus

$$(x_1, x_2, \cdots, x_{i-1}, \bar{x}_i, x_{i+1}, \cdots, x_{n-1}) \in M_{j,k}$$

and it has been shown that $M_{j,k}$ is linearly closed.

The induction assumption may now be applied to the sequence $\{M_{j,k}\}$ and it is found that, for some integers j_1, k_1 the set M_{j_1,k_1} contains a subinterval of U_{n-1}. Let J_1 be a closed subinterval of U_{n-1} contained in M_{j_1,k_1}. Thus, for every point $\mathbf{x}_{n-1}$ of the interval J_1 there is an interval $I=I(\mathbf{x}_{n-1})\subset U$ such that

$$|I(\mathbf{x}_{n-1})| \geqq 1/j_1$$

and

$$x_n \in I(\mathbf{x}_{n-1}) \text{ implies } (\mathbf{x}_{n-1}, x_n) \in S_{k_1}.$$

Now let K_j, for $j=1, 2, \cdots, 2j_1$, denote the interval

$$K_j = ((j-1)/(2j_1), j/(2j_1)) \subset U$$

of length $1/(2j_1)$. Then let N_j, for $j=1, 2, \cdots, 2j_1$ be the set

$$N_j \equiv [\mathbf{x}_{n-1} \in J_1;\ x_n \in K_j \text{ implies } (\mathbf{x}_{n-1}, x_n) \in S_{k_1}].$$

Then it is very easily seen that N_j is linearly closed. On the other hand, if $\mathbf{x}_{n-1}\in J_1$ then $I(\mathbf{x}_{n-1})$, being an interval of length at least $1/j_1$ must contain some one of the intervals K_j. Then $\mathbf{x}_{n-1}\in N_j$ for some j, that is, $\mathbf{x}_{n-1}\in\sum N_j$. Thus $\sum N_j\supset J_1$ and so $\sum N_j$ has property Δ by Baire's theorem. Thus, again applying the induction hypothesis, some N_{j_2} contains a subinterval $J_2\subset J_1\subset U_{n-1}$. Then the subinterval $J_2\times K_{j_2}$ of U_n is contained in S_k. This completes the proof of Lemma 2.

Another property of subsets of U_n which is connected with property Δ will be useful later and is defined now.

Property Γ. Again the definition is inductive for sets in U_n. A set $S\subset U_1$ has property Γ if it is residual in U_1. A set $S\subset U_n$ has property Γ if there is a set $R\subset U_{n-1}$ such that

$$R \text{ has property } \Gamma \text{ in } U_{n-1}$$

and such that if $\mathbf{x}_{n-1}\in R$ then the set

$$[x_n \in U;\ (\mathbf{x}_{n-1}, x_n) \in S]$$

is residual in U.

Clearly a set with property Γ has also property Δ, not conversely. However, there is a more striking connection expressed by the following:

LEMMA 3. *A set $S \subset U_n$ has property Γ if and only if $U_n - S$ does not have property Δ.*

Proof. This can be proved by a straightforward induction. (If $n=1$ the statement reduces to the definition of a residual set.) The details of the proof suggest themselves readily.

Projection. If $S \subset U_n$ then, for each $i=1, 2, \cdots, n$, the set of all

$$(x_1, x_2, \cdots, x_{i-1}, x_{i+1}, \cdots, x_n) \in U_{n-1}$$

such that there exists an $x_i \in U$ for which

$$(x_1, x_2, \cdots, x_{i-1}, x_i, x_{i+1}, \cdots, x_n) \in S$$

is called the projection of S on U^i_{n-1}.

The last set-theoretic lemma that will be needed involves a connection between the property Δ and the dimension of a subset of U_n.

LEMMA 4. *Let $S \subset U_n$. Suppose that the projection of S on U^i_{n-1} for each $i=1, 2, \cdots, n$, fails to have property Δ. Then S has dimension at most $n-2$.*

Proof. Again this can be shown by a straightforward induction which it does not seem necessary to present in detail. It should be mentioned that one uses Lemma 3 and two well known facts about residual sets, namely that the product of two residual sets is residual and that a residual set is dense.

3. **Further preliminaries.** Throughout this section let $f(x_1, x_2, \cdots, x_n) = f(\mathbf{x}_{n-1}, x_n) = f(\mathbf{x}_n)$ be defined in U_n and be a continuous function of each x_i for fixed $x_1, x_2, \cdots, x_{i-1}, x_{i+1}, \cdots, x_n$. In particular $f(\mathbf{x}_{n-1}, x_n)$ is a continuous function of x_n in U.

Unicontinuous. For convenience a function f satisfying the above conditions will be referred to as unicontinuous.

$\delta_\epsilon(\mathbf{x}_{n-1})$. For fixed $\mathbf{x}_{n-1} \in U_{n-1}$, f is a uniformly continuous function of x_n, that is, for every $\epsilon > 0$ there is a δ_ϵ such that

$$|x_n - \bar{x}_n| \leqq \delta_\epsilon$$

implies

$$|f(\mathbf{x}_{n-1}, x_n) - f(\mathbf{x}_{n-1}, \bar{x}_n)| \leqq \epsilon.$$

It is apparent that for each $\epsilon > 0$ and for fixed $\mathbf{x}_{n-1}$, there is a greatest such δ_ϵ. This greatest δ_ϵ will be denoted by $\delta_\epsilon(\mathbf{x}_{n-1})$.

LEMMA 5. *For each $\epsilon > 0$, $\eta > 0$ the set of points*

$$[\mathbf{x}_{n-1} \in U_{n-1}; \delta_\epsilon(\mathbf{x}_{n-1}) \geqq \eta]$$

is linearly closed.

Proof. This statement, which expresses the upper semi-continuity of

$\delta_\epsilon(x_1, x_2, \cdots, x_{n-1})$ with respect to each of $x_1, \cdots, x_{n-1}$, is reasonably well known and in any case is trivial.

$\Omega(f, S)$. If S is a subset of U_n then $\Omega(f, S)$ will denote

$$\Omega(f, S) \equiv \text{l.u.b. } [f(\mathbf{x}_n); \mathbf{x}_n \in S] - \text{g.l.b. } [f(\mathbf{x}_n); \mathbf{x}_n \in S].$$

Clearly $S \subset T$ implies $0 \leqq \Omega(f, S) \leqq \Omega(f, T)$.

$\Omega(f, \mathbf{x}_n)$. If $\mathbf{x}_n \in U_n$ then $\Omega(f, \mathbf{x}_n)$ will denote

$$\Omega(f, \mathbf{x}_n) \equiv \text{g.l.b. } [\Omega(f, S); S \text{ is open and } \mathbf{x}_n \in S].$$

LEMMA 6. *If $\mathbf{x}_n \in U_n$ and S_n is the open cube of side $1/n$ centered at $\mathbf{x}_n$ then*

$$\Omega(f, \mathbf{x}_n) = \lim_{n\to\infty} \Omega(f, S_n).$$

LEMMA 7. *The function f is continuous at $\mathbf{x}_n$ if and only if $\Omega(f, \mathbf{x}_n) = 0$.*

LEMMA 8. *For each $\eta > 0$ the set*

$$[\mathbf{x}_n \in U; \Omega(f, \mathbf{x}_n) \geqq \eta]$$

is closed.

These three well known facts are stated only for reference.

4. **Functions of two variables.** This section will be devoted to the consideration of a unicontinuous function $f(x_1, x_2)$ defined in U_2.

THEOREM 1. *Let $f(x_1, x_2) = f(\mathbf{x}_2)$ be unicontinuous in U_2. Then for each $\eta > 0$ the set*

$$D_\eta \equiv [\mathbf{x}_2 \in U_2; \Omega(f, \mathbf{x}_2) \geqq \eta]$$

has a nowhere dense projection on U_1^1 and U_1^2.

Proof. In view of the symmetry of the assumptions it is sufficient to prove that D_η has a nowhere dense projection on U_1^1.

Let ϵ be fixed in $0 < \epsilon < \eta/4$. Then let

$$S_j \equiv [x_1 \in U_1; \delta_\epsilon(x_1) \geqq 1/j], \qquad j = 1, 2, \cdots,$$

where $\delta_\epsilon(x_1) = \delta_\epsilon(\mathbf{x}_{2-1})$ was defined in §2. Then S_j is closed, by Lemma 5. Also $\sum S_j = U$ is of second category by Baire's theorem. Hence, by Lemma 1, there is an integer k and an interval $I \subset U$ such that $I \subset S_k$. Thus

$$x_1 \in I \text{ implies } \delta_\epsilon(x_1) \geqq 1/k. \tag{1}$$

It will now be shown that if

$$(x_1, x_2) = \mathbf{x}_2 \in I \times U$$

then

$$\Omega(f, \mathbf{x}_2) \leqq 4\epsilon.$$

To this end let $\bar{x}_1 \in I$ and $\bar{x}_2 \in U$. Then, since $f(x_1, \bar{x}_2)$ is a continuous function of x_1, there is a $\delta > 0$ such that if $x_1 \in (\bar{x}_1 - \delta, \bar{x}_1 + \delta)$ then

$$| f(x_1, \bar{x}_2) - f(\bar{x}_1, \bar{x}_2) | \leqq \epsilon.$$

Choose $\delta > 0$ so small that $(\bar{x}_1 - \delta, \bar{x}_1 + \delta) \subset I$, as may be done since I is open. Then, by (1), for each $x_1 \in (\bar{x}_1 - \delta, \bar{x}_1 + \delta) \subset I$

$$| f(x_1, x_2) - f(x_1, \bar{x}_2) | \leqq \epsilon$$

provided

$$| x_2 - \bar{x}_2 | \leqq 1/k \leqq \delta_\epsilon(x_1).$$

Combination of the last two inequalities shows that the function value at any point (x_1, x_2) in the rectangle

$$J \equiv (\bar{x}_1 - \delta, \bar{x}_1 + \delta) \times (\bar{x}_2 - 1/k, \bar{x}_2 + 1/k)$$

differs from the function value at (x_1, x_2) by at most 2ϵ. Thus

$$\Omega(f, J) \leqq 4\epsilon.$$

(I do not pause over the modifications necessary at the boundary of U_2.) Thus, a fortiori,

$$\Omega(f, (\bar{x}_1, \bar{x}_2)) \leqq 4\epsilon < \eta.$$

But $(\bar{x}_1, \bar{x}_2)$ was any point $I \times U$.

It has been shown that U_2 contains a strip $I \times U$ which contains no point of D_η. But the argument given was equally applicable to any substrip of U_2. This completes the proof of Theorem 1.

THEOREM 2. *Let $f(x_1, x_2)$ be unicontinuous in U_2. Let D denote the set of points in U_2 where $f(x_1, x_2)$ is discontinuous. Then D is an F_σ-set and the projection of D on U_1^1 (U_1^2) is of first category.*

Proof. If $D_{1/n}$ is the set

$$D_{1/n} \equiv [x_2 \in U_2; \Omega(f, x_2) \geqq 1/n]$$

then $D_{1/n}$ is closed, by Lemma 8. Also $D_{1/n}$ has a nowhere dense projection on U_1^1 by Theorem 1. Since $\sum D_{1/n} = D$, by Lemma 7, the proof is complete.

The next statement is the one which was called (A) in the introduction. It is clearly a restatement of Theorem 2.

THEOREM 3. *Let $f(x_1, x_2)$ be unicontinuous in U_2. Then there is a residual set of lines parallel to each axis consisting entirely of points where $f(x_1, x_2)$ is continuous.*

THEOREM 4. *The set of discontinuities of a unicontinuous function in U_2 has dimension at most zero.*

Proof. This is immediate from Theorem 3 in view of the fact that a residual set is dense. Of course it follows from Lemma 4 also.

5. **Functions of many variables.** This section will be devoted to a proof of the appropriate generalization of the results of the last section for a unicontinuous function on U_n.

THEOREM 5. *Let $f(x_1, \cdots, x_n) = f(\mathbf{x}_n)$ be unicontinuous in U_n. Then for each $\eta > 0$ the set*

$$D_\eta \equiv [\mathbf{x}_n \in U_n;\ \Omega(f, \mathbf{x}_n) \geqq \eta]$$

has a nowhere dense projection on U_{n-1}^i for each $i = 1, 2, \cdots, n$.

Proof. Clearly it is sufficient to prove that the projection of D_η on U_{n-1}^n is nowhere dense.

Let ϵ be fixed in $0 < \epsilon < \eta/8$. Then let

$$S_j \equiv [\mathbf{x}_{n-1} \in U_{n-1};\ \delta_\epsilon(\mathbf{x}_{n-1}) \geqq 1/j], \qquad j = 1, 2, \cdots.$$

Then S_j is linearly closed by Lemma 5. Also $\sum S_j = U_{n-1}$ has property Δ in U_{n-1} by Baire's theorem. Hence, by Lemma 2, there is an integer k and an interval $I_{n-1} \subset U_{n-1}$ such that $I_{n-1} \subset S_k$. Thus

$$\text{(1)} \qquad \mathbf{x}_{n-1} \in I_{n-1} \text{ implies } \delta_\epsilon(\mathbf{x}_{n-1}) \geqq 1/k.$$

Now suppose the theorem true for functions unicontinuous in U_{n-1}. For each fixed $x_n \in U$ let $f_{x_n}(\mathbf{x}_{n-1})$ denote the function $f(\mathbf{x}_{n-1}, x_n)$ of $\mathbf{x}_{n-1}$. Let $\{x_{n,h}\}$ be a dense sequence in U. Finally let

$$D_\epsilon^h \equiv [\mathbf{x}_{n-1} \in U_{n-1};\ \Omega(f_{x_{n,h}}, \mathbf{x}_{n-1}) \geqq \epsilon].$$

Then by the induction assumption D_ϵ^h has a nowhere dense projection on U_{n-2} and, a fortiori, D_ϵ^h is nowhere dense in U_{n-1}. Thus, by Baire's theorem,

$$\sum_h D_\epsilon^h \text{ does not contain } I_{n-1}.$$

Thus there is a point $\bar{\mathbf{x}}_{n-1} \in I_{n-1}$ such that $\bar{\mathbf{x}}_{n-1}$ is not in any D_ϵ^h, that is, so that

$$\text{(2)} \qquad \Omega(f_{x_{n,h}}, \bar{\mathbf{x}}_{n-1}) < \epsilon.$$

It will now be shown that for every $x_n \in U$,

$$\Omega(f, (\bar{\mathbf{x}}_{n-1}, x_n)) < 8\epsilon.$$

In fact let $\bar{x}_n \in U$ be fixed. Since $\{x_{n,h}\}$ is a dense sequence there is some h_1 such that

$$|\bar{x}_n - x_{n,h_1}| \leqq 1/2k.$$

But, since $\bar{\mathbf{x}}_{n-1} \in I_{n-1}$, this means, by (1),

$$|\bar{x}_n - x_{n,h_1}| \leqq \delta_\epsilon(\bar{\mathbf{x}}_{n-1})$$

so that

(3) $$|f(\bar{\mathbf{x}}_{n-1}, \bar{x}_n) - f(\bar{\mathbf{x}}_{n-1}, x_{n,h_1})| \leqq \epsilon,$$

by the definition of $\delta_\epsilon(\mathbf{x}_{n-1})$. Now, by (2) and the definition of $\Omega(f_{x_{n,h}}, \bar{\mathbf{x}}_{n-1})$, there is an open neighborhood $J_{n-1}^{h_1}$ of $\bar{\mathbf{x}}_{n-1}$ in U_{n-1} such that

$$\Omega(f_{x_{n,h_1}}, J_{n-1}^{h_1}) < 2\epsilon.$$

Thus

(4) $$\bar{\bar{\mathbf{x}}}_{n-1} \in J_{n-1}^{h_1} \text{ implies } |f(\bar{\mathbf{x}}_{n-1}, x_{n,h_1}) - f(\bar{\bar{\mathbf{x}}}_{n-1}, x_{n,h_1})| < 2\epsilon.$$

Now suppose $J_{n-1}^{h_1}$ has been chosen so "small" that $J_{n-1}^{h_1} \subset I_{n-1}$ which is clearly permissible. Then

$$\bar{\bar{\mathbf{x}}}_{n-1} \in J_{n-1}^{h_1} \subset I_{n-1} \text{ implies } \delta_\epsilon(\bar{\bar{\mathbf{x}}}_{n-1}) \geqq 1/k.$$

In other words

(5) $$|\bar{\bar{x}}_n - x_{n,h_1}| \leqq 1/k \text{ implies } |f(\bar{\bar{\mathbf{x}}}_{n-1}, x_{n,h_1}) - f(\bar{\bar{\mathbf{x}}}_{n-1}, \bar{\bar{x}}_n)| \leqq \epsilon.$$

Combination of (3), (4), (5) gives

(6) $$|f(\bar{\mathbf{x}}_{n-1}, \bar{x}_n) - f(\bar{\bar{\mathbf{x}}}_{n-1}, \bar{\bar{x}}_n)| < 4\epsilon$$

whenever

(7) $$\bar{\bar{\mathbf{x}}}_{n-1} \in J_{n-1}^{h_1} \text{ and } |\bar{\bar{x}}_n - x_{n,h_1}| \leqq 1/k$$

where h_1 was determined so that

$$|x_{n,h_1} - \bar{x}_n| \leqq 1/2k.$$

In particular if $|\bar{\bar{x}}_n - \bar{x}_n| \leqq 1/2k$ the second requirement (6) is automatic. Thus for any point $(\mathbf{x}_{n-1}, \bar{\bar{x}}_n)$ of the neighborhood

$$I_n \equiv J_{n-1}^{h_1} \times (\bar{x}_n - 1/2k, \bar{x}_n + 1/2k)$$

of $(\bar{\mathbf{x}}_{n-1}, \bar{x}_n)$ the inequality (6) is valid, Then,

$$\Omega(f, I_n) < 8\epsilon$$

and, a fortiori,

$$\Omega(f, \bar{x}_n) < 8\epsilon.$$

It has now been shown that the set

(8) $$[\mathbf{x}_n \in U_n; \Omega(f, \mathbf{x}_n) < 8\epsilon]$$

contains the line $\bar{\mathbf{x}}_{n-1} \times U$. But (8) is an open set (cf. Lemma 8) in U_n and so (8) contains a strip $J_{n-1} \times U$ for some interval $J_{n-1} \subset U_{n-1}$.

The above argument gives the existence of a strip $J_{n-1}\times U$ in the complement of $D_\eta \subset D_{8\epsilon}$ (since $8\epsilon<\eta$). But this argument was equally applicable to any substrip of U_n. This completes the proof of Theorem 5.

Exactly as in the case of two variables the following is an immediate consequence.

THEOREM 6. *Let $f(x_n)$ be unicontinuous in U_n. Let D denote the set of points in U_n where $f(x_n)$ is discontinuous. Then D is an F_σ-set and the projection of D on U^i_{n-1} is (for each $i=1, 2, \cdots, n$) of first category.*

This result, which is the natural extension of Theorem 2, does not obviously imply the desired generalization of Theorem 4. However, in view of Lemma 2, Theorem 6 is equivalent to the following:

THEOREM 7. *Let $f(x_n)$ be unicontinuous in U_n. Let D denote the set of points in U_n where $f(x_n)$ is discontinuous. Then the projection of D on U^i_{n-1} (for each $i=1, 2, \cdots, n$) does not have property Δ.*

Proof. In view of Lemma 1, an F_σ-set is of second category if and only if it contains an interval. In view of Lemma 2 and the fact that a closed set is linearly closed, an F_σ-set has property Δ if and only if it contains an interval. Thus, for F_σ-sets, first category is equivalent to the negation of property Δ. This shows the equivalence of Theorems 6 and 7.

Of course Theorem 7 may be stated in a positive fashion similar to Theorem 3. It is this statement which was displayed, for the case $n=3$, in the introduction. Finally

THEOREM 8. *The set of discontinuities of a unicontinuous function in U_n has dimension at most $n-2$.*

Proof. This is immediate from Theorem 7 and Lemma 4.

6. **An example.** In this section it will be shown that Theorem 6 describes the possible sets D completely. This is done by proving the following.

THEOREM 9. *Let D be any F_σ-set in U_n such that the projection of D on U^i_{n-1} (for each $i=1, 2, \cdots, n$) is of first category. Then there exists a unicontinuous function on U_n for which D is the set of discontinuity points.*

The example which proves Theorem 9 will be constructed with the help of certain auxiliary functions whose existence is demonstrated first.

LEMMA 9. *Let D be any closed set in U_n such that the projection of D on U^i_{n-1} (for each $i=1, 2, \cdots, n$) is nowhere dense. Then there exists a function $f=f_D$ on U_n such that*

(a) $0\leq f_D\leq 1$;

(b) *f_D is unicontinuous on U_n;*

(c) $\Omega(f_D, x_n)=0$, *if* $x_n\in U_n-D$;

(d) $\Omega(f_D, x_n)=1$, *if* $x_n\in D$.

Proof. By an oriented closed cube K in U_n will be understood a closed n-cube with faces parallel to the coordinate hyperplanes. With each oriented cube $K \subset U_n$ let there be associated a function g_K with the following properties:

(1) g_K is defined and continuous on K;

(2) $g_K = 0$ on the boundary of K;

(3) $0 \leqq g_K \leqq 1$;

(4) $g_K = 1$ at the midpoint of K.

For example $g_K(x_n)$ might be chosen proportional to the distance from x_n to the boundary of K.

Let a line, parallel to the ith coordinate axis of U_n ($i = 1, 2, \cdots, n$) be called an i-grid line if it contains a point of D. Let D_i^* be the set of all points which lie on some i-grid line. Then D_i^* is clearly a closed set since D is closed. Also D_i^* has the same projection on U_{n-1}^i as D, hence D_i^* is nowhere dense in U_n.

Now let $D^* \equiv \sum_{i=1}^n D_i^*$. Then D^* is a nowhere dense closed set in U_n. Let $C^* = U_n - D^*$, so that C^* is an open set dense in U_n. It is well known that any open set in U_n is the sum of a countable number of nonoverlapping oriented closed cubes. Thus $C^* = \sum K_i$ where K_i is an oriented closed cube in U_n and K_i and K_j have at most boundary points in common.

Notice that every point of D is a limit point of midpoints of the K_i. In fact $\sum K_i = C^*$ is dense in U_n and so each point of D is a limit point of points in some collection of K_i. But any finite collection of K_i form a closed set disjoint from D^* and, a fortiori, disjoint from D. Thus each point of D is a limit point of points from distinct K_i. But the diameters of any infinite collection of K_i obviously must approach zero since $\sum K_i \subset U_n$. Thus each point of D is a limit point of midpoints of K_i.

Now let $\{P_j\}$ be a sequence of points in D such that every point of D is either a point P_j or a limit point of points P_j. The existence of such a sequence is quite obvious. Let a subsequence $\{K_{m,j}\}$ of the cubes K_i be chosen in such a way that the midpoint of $K_{m,j}$ is at a distance at most $1/m$ from P_j. Tne fact that $\{K_{m,j}\}$ exists is clear from the preceding paragraph. It is clear that by proceeding inductively with respect to $m+j$ the $K_{m,j}$ might be chosen as all distinct, but this is not essential.

Finally let $f = f_D$ be defined in U_n by

$$(5)\quad f_D(x_n) \equiv \begin{cases} g_{K_{m,j}}(x_n) & \text{if } x_n \in K_{m,j}; \quad j = 1, 2, \cdots; m = j, j+1, \cdots, \\ 0 & \text{if } x_n \in U_n - \sum_{m \geqq j} K_{m,j}. \end{cases}$$

It is to be shown that this function f_D satisfies the requirements (a)–(d) of Lemma 9.

First of all (a) is obvious from (5) and (3).

In order to prove (d) let P be a point in D. Then either P is a point P_j or P

is a limit point of such points. In the first case $P=P_j$ then P is a limit point of the midpoints of the cubes $K_{m,j}$ for $m=j, j+1, \cdots$. Thus P is a limit point of points where $f_D=1$ (in view of (5), (4)). But also P is a limit point of boundary points of $K_{m,j}$ where $f_D=0$. Thus $\Omega(f_D, P)=1$. In the other case, that P is the limit of some sequence $\{P_{j_i}\}$, then P is also the limit of the midpoints of K_{j_i,j_i} and again $\Omega(f_D, P)=1$. This establishes (d).

To prove (c), let P be a point where $\Omega(f_D, P)>0$. Then P is not an interior point of any cube $K_{m,j}$ by (1). Thus $f_D(P)=0$ (and in fact P is a limit point of points where $f_D=0$). But, since $\Omega(f_D, P)=\delta>0$, P must also be a limit point of points where $f_D>\delta/2$ and hence a limit point of points in some subsequence $\{K_{m_ij_i}\}$ of $\{K_{m,j}\}$. This subsequence cannot contain only a finite number of distinct cubes since, if it did P would clearly be an interior point of some $K_{m,j}$. Thus P is the limit point of midpoints of an infinite sequence $\{K_{m_i,j_i}\}$ of distinct cubes. But each such midpoint is at a distance at most $1/m_i$ from P_{j_i}. Thus P is a point P_{j_i} or a limit point of such points and, in either case, $P\in D$, since D is closed.

Finally, to prove (b), the value of f_D along lines parallel to the axes must be considered. If such a line is a grid line then it is contained entirely in D^* and $f_D\equiv 0$ by (5). If such a line is not a grid line it is contained entirely in U_n-D so that f_D is continuous at every point by (c) and, a fortiori, is continuous along the given line.

The proof of Lemma 9 is now complete and it is now very easy to prove Theorem 9.

Proof of Theorem 9. Let D be any F_σ-set in U_n having a first category projection on U_{n-1}^i. Then, by the definition of an F_σ-set, $D=\sum D_i$, where D_i is closed. But D_i has a nowhere dense projection on each U_{n-1}^i. (In fact if a closed set is dense in some interval it contains the interval and so is not of first category by Baire's theorem.) Thus each D_i satisfies the requirement of the set D in Lemma 9. Let f_{D_i} be a function associated with D_i satisfying (a)–(d) of Lemma 9. Then let

$$f=\sum_{i=1}^{\infty} f_{D_i}/3^i.$$

Then the series defining f is uniformly convergent by (a). Thus f is unicontinuous since all f_{D_i} are unicontinuous. Also f is continuous on $U_n-D=U_n-\sum D_i$ since each f_{D_j} is continuous on $U_n-D_j\supset U_n-\sum D_i$. Finally, at any point P of D, f is discontinuous since the convergence factors $1/3^i$ were chosen sufficiently rapid that "cancellation" of the discontinuities is impossible in view of (d).

7. The consequence of stronger assumptions. This section and the following one are devoted to an investigation of the consequence of assuming stronger smoothness than simply continuity parallel to the axes. In order

to state the precise requirement that will be considered in this section, a definition is needed first.

Let $f(\mathbf{x}_n)$ be a unicontinuous function in U_n, so that $f(\mathbf{x}_{n-1}, x_n)$ is a uniformly continuous function of x_n in U. Let

$$\omega(\mathbf{x}_{n-1}, \delta) = \max_{|x_n - x_n'| \leq \delta} | f(\mathbf{x}_{n-1}, x_n') - f(\mathbf{x}_{n-1}, x_n'') |.$$

Then $\omega(\mathbf{x}_{n-1}, \delta)$ is, for each fixed $\mathbf{x}_{n-1}$, a monotone non-decreasing function of δ in $0<\delta\leqq 1$ with $\omega(\mathbf{x}_{n-1}, 0+0)=0$.

Condition S: The function $f(\mathbf{x}_{n-1}, x_n)$ of x_n is said to satisfy condition S if there exists a sequence $\{\omega_m(\delta)\}$ of functions $\omega_m(\delta)$ defined for $0<\delta\leqq 1$ such that

$$\omega_m(0 + 0) = 0$$

and, for every $\mathbf{x}_{n-1} \in U_{n-1}$ there is an integer m such that

$$\omega(x_{n-1}, \delta) \leqq \omega_m(\delta), \qquad 0 < \delta \leqq 1.$$

For example if, for each fixed $\mathbf{x}_{n-1}$, f satisfies a Lipschitz condition of order one, such a sequence is provided by $\omega_m(\delta)=m\cdot\delta$. If it is known only that for each $\mathbf{x}_{n-1}$, f satisfies a Lipschitz condition of some positive order then $\omega_m(\delta)=m\cdot\delta^{1/m}$ is a sequence of the required type. By considering such sequences as $\omega_m(\delta)=1/|\log \delta|^{1/m}$ it is seen that condition S is very weak. However the theorem to be proved next, taken in conjunction with the example of the preceding section, implicitly shows that condition S is not always satisfied.

THEOREM 10. *Let $f(\mathbf{x}_n)$ be a unicontinuous function on U_n which satisfies condition* S *with respect to any one of its variables as x_n. Let D denote the set of points in U_n where $f(\mathbf{x}_n)$ is discontinuous. Then D is an F_σ-set and the projection of D on U_{n-1}^i is (for each $i=1, 2, \cdots, n$) nowhere dense.*

Proof. The proof is by induction based on the completely trivial case $n=1$. Suppose then that the theorem is true for the case $n-1$.

Now let

$$S_j = [\mathbf{x}_{n-1};\ \omega(\mathbf{x}_{n-1}, \delta) \leqq \omega_j(\delta)].$$

Then S_j is easily seen to be linearly closed. Also $\sum S_j = U_{n-1}$ in view of condition S. Thus Lemma 2 can be applied to the sequence $\{S_j\}$ and one finds that there is an integer k and a subinterval $I_{n-1} \subset U_{n-1}$ such that $I_{n-1} \subset S_k \subset U_{n-1}$.

Let $\bar{x}_n \in U$ be fixed. Then $f(\mathbf{x}_{n-1}, \bar{x}_n)$, considered as a function of $\mathbf{x}_{n-1}$, has, by the induction hypothesis, discontinuities at a set of $\mathbf{x}_{n-1}$ with a nowhere dense projection and, a fortiori, nowhere dense in U_{n-1}. Thus there is an interval $J_{n-1} \subset I_{n-1}$ where $f(\mathbf{x}_{n-1}, \bar{x}_n)$ is a continuous function of $\mathbf{x}_{n-1}$. Thus if $\bar{\mathbf{x}}_{n-1} \in J_{n-1}$ then for every $\epsilon>0$ there is a δ_ϵ such that

$$\text{(1)} \qquad \begin{gathered} | \bar{\mathbf{x}}_{n-1} - \mathbf{x}_{n-1} | < \delta_\epsilon \text{ implies } \mathbf{x}_{n-1} \in J_{n-1} \text{ and} \\ | f(\bar{\mathbf{x}}_{n-1}, \bar{x}_n) - f(\mathbf{x}_{n-1}, \bar{x}_n) | < \epsilon/2. \end{gathered}$$

Now let $\bar{\delta}>0$ be chosen so that

$$\omega_k(\bar{\delta}) \leqq \epsilon/2$$

which is possible since $\omega_k(0+0)=0$. Then, for any $x_{n-1}\in J_{n-1}\subset I_{n-1}\subset S_k$,

$$\omega(x_{n-1}, \bar{\delta}) \leqq \omega_k(\bar{\delta}) \leqq \epsilon/2.$$

Thus, by the definition of $\omega(x_{n-1}, \delta)$,

$$(2)\qquad | \bar{x}_n - x_n | < \bar{\delta} \text{ implies } | f(x_{n-1}, \bar{x}_n) - f(x_{n-1}, x_n) | \leqq \epsilon/2.$$

From (1) and (2), one has

$$| f(\bar{x}_{n-1}, \bar{x}_n) - f(x_{n-1}, x_n) | \leqq \epsilon$$

if $|\bar{x}_{n-1}-x_{n-1}|<\delta_\epsilon$ and $|\bar{x}_n-x_n|<\bar{\delta}$. Hence f is continuous at $(\bar{x}_{n-1}, \bar{x}_n)$. But $(\bar{x}_{n-1}, \bar{x}_n)$ was arbitrary in $J_{n-1}\times U$.

It has been shown that U_n contains a strip $J_{n-1}\times U$ consisting entirely of continuity points of f. But the argument leading to the existence of this strip is equally applicable to any substrip of U_n. Thus the discontinuity points of f have a nowhere dense projection on U_{n-1}^n. The symmetry of the assumptions shows that the proof is complete.

8. **Another example.** In this section it will be shown that the conclusion of Theorem 10 cannot be strengthened (as far as a restriction on D is concerned) even if the assumptions are strengthened to the point of requiring the existence of all derivatives along any line parallel to an axis. The precise statement of the theorem to be established follows.

THEOREM 11. *Let D be any F_σ-set in U_n such that the projection of D on U_{n-1}^i (for each $i=1, 2, \cdots, n$) is nowhere dense. Then there exists a function $f(x_n)$ on U_n such that $f(x_n)$, considered as a function of any one of the variables (for fixed values of the remaining $n-1$ variables) has all derivatives and such that D is the set of discontinuity points of $f(x_n)$.*

Proof. The construction of the example which proves Theorem 11 is similar to the construction in the proof of Lemma 9 and will not be given in quite so much detail.

First let $\bar{D}$ be the closure of D. Then $\bar{D}$ has a nowhere dense projection on each U_{n-1}^i. As in the proof of Lemma 9 let $\bar{D}_i^*$ be the set of points lying on some "i-grid line," that is, some line parallel to the x_i-axis containing a point of $\bar{D}$. As before, let $\bar{D}^*=\sum_{i=1}^n \bar{D}_i^*$. Finally let $\bar{C}^*=U_n-\bar{D}^*$, so that $\bar{C}^*$ is a dense open set in U_n. Now $\bar{C}^*=\sum K_i$ where K_i is a closed oriented cube and K_i and K_j have at most boundary points in common.

With any closed oriented cube $K\subset U_n$ let there be associated a function g_K which, in addition to satisfying the requirements (1)–(4) of §6, is infinitely differentiable along any line parallel to an axis, with all derivatives vanishing at the boundary of K. (Of course, at the boundary of K these derivatives are

all one-sided derivatives calculated with respect to the interior of K.) For example, if K is the unit cube U_n then g_K might be chosen as

$$g_K(x_1, x_2, \cdots, x_n) = \exp(8n)\prod_{i=1}^{n} \exp(-1/x_i^2)\cdot\exp(-1/(x_i-1)^2).$$

Now let the F_σ-set D be written as $D=\sum D_k$ where D_k is closed. For each k let $\{P_{k,j}\}$ be a sequence of points in D_k such that every point of D_k is either a point $P_{k,j}$ or a limit point of such points. Also let $\{K_{m,k,j}\}$ be a subsequence of the cubes K_i such that the midpoint of $K_{m,k,j}$ is at a distance at most $1/m$ from $P_{k,j}$. This time it is important to choose the $K_{m,k,j}$ all distinct.

Now a function $f(\mathbf{x}_n)$ satisfying the requirements of Theorem 11 can be defined by

$$f(\mathbf{x}_n) = \begin{cases} (1/k)g_{K_{m,k,j}}(\mathbf{x}_n) & \text{if } \mathbf{x}_n \in K_{m,k,j}, \qquad m \geqq \max(k, j); \\ 0 & \text{if } \mathbf{x}_n \in U_n - \sum_{m\geqq k,j} K_{m,k,j}. \end{cases}$$

The fact that this function is discontinuous on D and continuous on $U_n - D$ is verified very much as it was shown that the function f_D defined in the proof of Lemma 9 had D as its set of discontinuities. In the present case D_K turns out to be the set of points P where $\Omega(f, P) \geqq 1/k$. To see that f is infinitely differentiable along any line parallel to a coordinate axis, notice that such a line lies either in $\overline{D}^*$ or in $\overline{C}^*$. In the first case $f\equiv 0$. In the second case the given line is at a finite distance $\delta>0$ from the closed set $\overline{D}^*$ and, a fortiori, at a distance at least $\delta>0$ from any point of $D\subset\overline{D}^*$. Thus at most a finite number of the cubes $K_{m,k,j}$, with $m\geqq\max(k, j)$, are intersected by the given line. Thus along the given line f is zero save for a finite number of nonoverlapping intervals in which it is modified by inserting an infinitely differentiable piece with all derivatives vanishing at the end points. This completes the proof.

THE JOHNS HOPKINS UNIVERSITY,
BALTIMORE, MD.

THE GAUSS-BONNET THEOREM FOR RIEMANNIAN POLYHEDRA

BY

CARL B. ALLENDOERFER AND ANDRÉ WEIL

TABLE OF CONTENTS

1. Introduction. The classical Gauss-Bonnet theorem expresses the "curvatura integra," that is, the integral of the Gaussian curvature, of a curved polygon in terms of the angles of the polygon and of the geodesic curvatures of its edges. An important consequence is that the "curvatura integra" of a closed surface (or more generally of a closed two-dimensional Riemannian manifold) is a topological invariant, namely (except for a constant factor) the Euler characteristic.

One of us(1) and W. Fenchel(2) have independently generalized the latter result to manifolds of higher dimension which can be imbedded in some Euclidean space. For such manifolds, they proved a theorem which we shall show to hold without any restriction, and which may be stated as follows:

THEOREM I. *Let M^n be a closed Riemannian manifold of dimension n, with the Euler-Poincaré characteristic χ; let $dv(z)$ be the Riemannian volume-element at the point with local coordinates z^μ $(1 \leqq \mu \leqq n)$; let $g_{\mu\nu}$ be the metric tensor, $g = |g_{\mu\nu}|$ its determinant, $R_{\mu_1\mu_2\nu_1\nu_2}$ the Riemannian curvature tensor at the same point; and define the invariant scalar $\Psi(z)$ by:*

$$\Psi(z) = (2\pi)^{-n/2} \cdot \frac{1}{2^n (n/2)!} \sum_{\mu,\nu} \frac{\epsilon^{(\mu)}\epsilon^{(\nu)}}{g} \cdot R_{\mu_1\mu_2\nu_1\nu_2} R_{\mu_3\mu_4\nu_3\nu_4} \cdots R_{\mu_{n-1}\mu_n\nu_{n-1}\nu_n} \quad \textit{for } n \textit{ even} \tag{1}$$

$$\Psi(z) = 0 \quad \textit{for } n \textit{ odd.}$$

Presented to the Society, December 30, 1941 under the title *A general proof of the Gauss-Bonnet theorem*; received by the editors April 23, 1942.

(1) C. B. Allendoerfer, *The Euler number of a Riemann manifold*, Amer. J. Math. vol. 62 (1940) p. 243.

(2) W. Fenchel, *On total curvatures of Riemannian manifolds.* (I), J. London Math. Soc. vol. 15 (1940) p. 15. The concluding words of this paper show that the author contemplated an extension of his method which was to give him "a formula of Gauss-Bonnet type." We do not know whether such an extension has been published, or even carried out.

Then:

$$\chi = \int_{M^n} \Psi(z) dv(z).$$

Here and throughout this paper a sign such as $\sum_{\mu,\nu}$ indicates summation over all indices μ_i, ν_i, these indices running *independently* over their whole range; and $\epsilon^{(\mu)}$ is the relative tensor $\epsilon^{\mu_1\mu_2\cdots\mu_n}$ defined by $\epsilon^{(\mu)}=+1$ if $(\mu_1, \mu_2, \cdots, \mu_n)$ is an even permutation of $(1, 2, \cdots . n)$, $\epsilon^{(\mu)}=-1$ if it is an odd permutation, and $\epsilon^{(\mu)}=0$ otherwise. Owing to the symmetry properties of the curvature tensor it is readily seen that each term in our sum occurs $2^n(n/2)!$ times or a multiple of that number; for that reason, in our arrangement of the numerical factor, the sign Σ is preceded by the inverse of that integer, so that the sum under Σ, together with the factor immediately in front of it, is (except for $1/g$) a polynomial in the R's with integer coefficients; similar remarks apply to the other formulae in this paper. On the other hand, it may be convenient, for geometric reasons, to define the curvature as $K=\omega_n/2\cdot\Psi(z)$, where ω_n is the surface-area of the unit-sphere S^n in R^{n+1}, so that the curvature is 1 for that sphere if n is even(3); Theorem I then gives $\int K dv(z)=\omega_n\cdot\chi/2$.

It does not seem to be known at present whether every closed Riemannian manifold can be imbedded in a Euclidean space. However, the possibility of local imbedding, at least in the analytic case, has been proved by E. Cartan(4), and this naturally suggests applying the same method of tubes, which was developed for closed imbedded manifolds in the above-mentioned paper(1), to the cells of a sufficiently fine subdivision of an arbitrary manifold. This gives a theorem on imbedded cells which is the n-dimensional analogue of the Gauss-Bonnet formula; the corresponding theorem for polyhedra will emerge as the main result of the present paper; except for details which will be filled in later, this can be stated as follows.

In a Riemannian manifold M^n, let M^r be a differentiable submanifold of dimension $r<n$; we assume that M^r is regularly imbedded in M^n, that is, taking local coordinates ζ^i on M^r and z^μ on M^n, that the matrix $\|\partial z^\mu/\partial\zeta^i\|$ is of rank r. We introduce the following tensors. First, we write:

$$\text{(2)} \qquad \mathrm{P}_{i_1i_2j_1j_2} = \sum_{\mu\nu} R_{\mu_1\mu_2\nu_1\nu_2}\cdot\frac{\partial z^{\mu_1}}{\partial\zeta^{i_1}}\cdot\frac{\partial z^{\mu_2}}{\partial\zeta^{i_2}}\cdot\frac{\partial z^{\nu_1}}{\partial\zeta^{j_1}}\cdot\frac{\partial z^{\nu_2}}{\partial\zeta^{j_2}},$$

(3) It will be noticed that for n even the numerical factor in $1/2\cdot\omega_n\cdot\Psi(z)$ as calculated from (1) has, owing to the value of $1/2\cdot\omega_n=2^{n/2}\cdot(2\pi)^{n/2}\cdot(n/2)!/n!$, a simple rational value.

(4) E. Cartan, *Sur la possibilité de plonger un espace riemannien donné dans un espace euclidien*, Annales de la Société Polonaise de Mathématique vol. 6 (1927) p. 1. This followed a paper by M. Janet under the same title, ibid. vol. 5 (1926) p. 38, where an incomplete proof of the same theorem is given; Janet's proof was completed by C. Burstin, *Ein Beitrag zum Problem der Einbettung der Riemannschen Räume in euklidischen Räumen*, Rec. Math. (Mat. Sbornik) N.S. vol. 38 (1931) p. 74.

those being the components of the curvature tensor of the imbedding manifold M^n in the directions which are tangent to M^r. Next, let x be a normal vector to M^r in M^n, with the covariant components x_ν; we write

$$(3)\qquad \Lambda_{ij}(x) = -\sum_\nu x_\nu\left[\frac{\partial^2 z^\nu}{\partial\zeta^i\partial\zeta^j} + \sum_{\lambda\mu}\begin{Bmatrix}\nu\\ \lambda\mu\end{Bmatrix}\frac{\partial z^\lambda}{\partial\zeta^i}\cdot\frac{\partial z^\mu}{\partial\zeta^j}\right].$$

where

$$\begin{Bmatrix}\nu\\ \lambda\mu\end{Bmatrix}$$

are the Christoffel symbols in M^n. The Λ's are linear combinations of the coefficients of the second fundamental form of M^r in M^n. We now introduce, for $0 \leqq 2f \leqq r$, the combinations

$$(4)\qquad \Phi_{r,f}(\zeta, x) = \frac{1}{2^{2f}\cdot f!(r-2f)!}\sum_{i,j}\frac{\epsilon^{(i)}\cdot\epsilon^{(j)}}{\gamma}$$
$$\cdot P_{i_1i_2j_1j_2}\cdots P_{i_{2f-1}i_{2f}j_{2f-1}j_{2f}}\cdot\Lambda_{i_{2f+1}j_{2f+1}}(x)\cdots\Lambda_{i_rj_r}(x)$$

where γ is the determinant of the metric tensor γ_{ij} on M^r. Let now S^{n-r-1} be the unit-sphere in the normal linear manifold $N^{n-r}(\zeta)$ to M^r at ζ; calling ξ an arbitrary point on that sphere, that is, an arbitrary unit-vector([5]), normal to M^r at ζ, we denote by $d\xi$ the area-element at ξ on S^{n-r-1}; and finally, we consider the expression([6])

$$(5)\qquad \Psi(\zeta, \xi\mid M^r) = \frac{\pi^{-n/2}\Gamma(n/2)}{2}\cdot\sum_{f=0}^{[r/2]}\frac{\Phi_{r,f}(\zeta,\xi)}{(n-2)(n-4)\cdots(n-2f)}\cdot d\xi,$$

which can be integrated over the whole or part of the sphere S^{n-r-1}.

Let now P^n be a Riemannian polyhedron, that is, a manifold with a boundary, the boundary consisting of polyhedra P^r_λ of lower dimensions (precise definitions will be given in §7); z^μ and ζ^i being local coordinates in P^n and P^r_λ, respectively, in the neighborhood of a point ζ of P^r_λ, we consider the set $\Gamma(\zeta)$ of all unit-vectors ξ at that point, with components ξ_μ such that $\sum_\mu\xi_\mu\cdot dz^\mu/ds \leqq 0$ when the derivatives dz^μ/ds are taken along any direction contained in the angle of P^n at ζ (for more details, see §§6–7). $\Gamma(\zeta)$ is found to be a spherical cell, bounded by "great spheres," on the unit-sphere S^{n-r-1} in the normal linear manifold to P^r_λ at ζ, and is what we call the "outer angle" of P^n at ζ.

([5]) We consistently (except for a short while in the proof of Lemma 8, §7) make no distinction between vectors and their end points, and therefore none between unit-vectors and points on the unit-sphere.

([6]) In view of the geometrical nature of the problem, one may suspect that the numerical coefficients in Ψ are connected with areas of spheres; and bringing out such connections may point the way to geometrical interpretations of our formulae. For instance, we have: $\pi^{-n/2}\cdot\Gamma(n/2)/(2\cdot 2^{2f}\cdot f!(r-2f)!(n-2)(n-4)\cdots(n-2f)) = 2/(\omega_{n-2f-1}\cdot\omega_{2f}\cdot(2f)!(r-2f)!2^f)$.

Our main theorem, which includes Theorem I as a particular case, expresses in terms of the above quantities the *inner characteristic* $\chi'(P^n)$ of P^n, that is, the Euler-Poincaré characteristic of the open complex consisting of all *inner* cells in an arbitrary simplicial or cellular subdivision of P^n; our methods would enable us to give a similar expression for the ordinary characteristic. The result is as follows:

THEOREM II. *P^n being a Riemannian polyhedron, with a boundary consisting of the polyhedra P_λ^r, we have*:

$$(-1)^n\chi'(P^n) = \int_{P^n} \Psi(z)dv(z) + \sum_{r=0}^{n-1}\sum_{\lambda} \int_{P_\lambda^r} dv(\zeta) \int_{\Gamma(\zeta)} \Psi(\zeta, \xi \mid P_\lambda^r).$$

It will be shown in §6 how the method of tubes, applied to an imbedded cell in a Euclidean space, leads directly to the formula in Theorem II for such a cell. Sections 2–3 give the necessary details on dual angles and outer angles, and contain the proof of the important additivity property for outer angles in affine space, which is stated in Theorem III; this may be considered as a theorem in spherical geometry, and is a wide generalization of some known results on polyhedra in R^3; it also includes some results of Poincaré on the angles of Euclidean and spherical polyhedra. Sections 4–5 are mainly devoted to the definition of the tube of a curved cell, and the investigation of its topological properties.

The proof of the main theorem then follows in §7, where it is shown how the additivity property for outer angles, proved in §2, implies an additivity property for the right-hand side in the formula in Theorem II; hence Theorem II is true for a polyhedron P^n if it is true for every polyhedron in a subdivision of P^n. In particular, it is true for an analytic cell because, by Cartan's theorem, every cell in a sufficiently fine subdivision of such a cell is imbeddable; by an elementary approximation theorem of H. Whitney, it is therefore true for an arbitrary cell. Hence it holds for every polyhedron which can be triangulated into cells; but it is known that every polyhedron can be so triangulated, and this completes the proof. Owing to the very unsatisfactory condition of our present knowledge of differentiable polyhedra, it has been found necessary to include, in §7, the proof of some very general lemmas on the subdivisions of such polyhedra; and the section concludes with some remarks about the validity of Theorem II for more general types of polyhedra than those we are dealing with.

2. **Dual angles in affine space.** It has often been observed that the word "angle" as used in elementary geometry is ambiguous, for it sometimes refers to a subset of the plane bounded by two rays and sometimes to what essentially is a 1-chain on the unit-circle. In order to preserve analogies with elementary geometry, we shall here use the word "angle" both for certain subsets of an affine vector-space R^n and for certain $(n-1)$-chains in the mani-

fold of directions from O in R^n; this will be done in such a way that no confusion may arise. Even in affine space we shall adopt the unit-sphere S^{n-1}, that is, the surface $\sum_\mu (x^\mu)^2=1$, as a convenient homeomorphic image of the manifold of directions from O in R^n; in the present section, any other such image could be used just as well to the same purpose.

In this section, R^n will denote an affine n-dimensional space over the field of real numbers. Assuming that a basis has been chosen in R^n once for all, we denote by x^μ $(1\leqq\mu\leqq n)$ the components of a vector x in R^n with respect to that basis. As functions of x, the n components x^μ are linear forms in R^n; and they constitute a basis for the vector-space $\bar{R}^n$ of all linear forms $(y, x)=\sum_\mu y_\mu\cdot x^\mu$ over R^n; the y_μ are then the components, with respect to that basis, of the form (y, x), or, as we may say for short, of the form y. We call $\bar{R}^n$ the dual space to R^n. We shall consider linear manifolds V^r in R^n, which, throughout §§2–3, should be understood to contain O; throughout this paper, the superscript, when used for a space or manifold, should be understood to indicate the dimension. To every V^r in R^n corresponds in $\bar{R}^n$ the dual manifold $\bar{V}^{n-r}$, consisting of all linear forms which vanish over V^r (this should not be confused with the dual space to V^r when the latter is considered as an affine space).

Convex angles in R^n may be defined in two ways, which may be considered as dual to each other: (a) a convex angle is the set of points x in R^n which satisfy a finite number of given inequalities $(b_\nu, x)\geqq 0$; (b) a convex angle is the set of points $x=\sum_\rho u_\rho\cdot a_\rho$, where the a_ρ are a finite number of given points, and the numbers u_ρ take all values greater than or equal to 0. It is well known that these two definitions are equivalent. Throughout this paper, all angles will be convex angles, and we shall often omit the word "convex."

A convex angle C is said to be of *dimension* r and of *type* s if r and s are the dimensions of the smallest linear manifold V^r such that $V^r\supset C$ and of the largest linear manifold V^s such that $C\supset V^s$; if $r=s$, the angle reduces to V^r and will be called *degenerate*; otherwise $r>s$. In the notation of angles, the superscript will usually denote the dimension and a Latin subscript the type of the angle whenever it is desirable to indicate either or both. A Greek subscript will be used to distinguish among angles of the same dimension and type.

Let C be an r-dimensional angle, contained in the linear manifold V^r; a point of C will be called an *inner* point if there is a neighborhood of that point in V^r which is contained in C; such points form a subset of C which is open with respect to the space V^r; if C is defined by the inequalities $(b_\nu, x)\geqq 0$, a point a in C will be an inner point if, and only if, all those of the forms (b_ν, x) which do not vanish on V^r are greater than 0 at a. For $r=0$, V^r and C both reduce to the point O, which is then considered as an inner point of C. The points of an angle which are not inner points constitute a set which is the union of angles of lower dimension; such points are limits of inner points.

LEMMA 1. *Let C be a convex angle of dimension r, with at least one point a in the open half-space $(b, x)>0$; then its intersection, D, with the closed half-space $(b, x)\geqq 0$ is a nondegenerate angle of the same dimension.*

For all points of C, in a sufficiently small neighborhood of a, will be in D; among those points there are inner points of C, forming an open set in the V^r which contains C, so that D is r-dimensional. Moreover, D contains a and not $-a$, and so cannot be degenerate.

Let C be a convex angle of dimension m; a finite set $\mathfrak{D}$ of distinct convex angles C_λ^r $(0\leqq r\leqq m;\ 1\leqq\lambda\leqq N_r)$ will be called a ***subdivision*** of C into convex angles whenever the two following conditions are fulfilled: (a) every point of C is an inner point of at least one C_λ^r in $\mathfrak{D}$; (b) if two angles C_λ^r, C_μ^s in $\mathfrak{D}$ are such that there is an inner point of C_λ^r which is contained in C_μ^s, then $C_\lambda^r\subset C_\mu^s$. From (b), it follows that no two distinct angles in $\mathfrak{D}$ can have an inner point in common. The angles in $\mathfrak{D}$ can be considered, in the usual way, as forming a combinatorial complex. A subdivision of an angle C is called ***degenerate*** if it contains a degenerate angle V^r of a dimension $r>0$; as O then is an inner point of V^r and is in all the angles of $\mathfrak{D}$, it follows that all those angles contain V^r and are of type at least r, as well as C itself. If $\mathfrak{D}$ is nondegenerate, it is easily shown to contain angles of all dimensions less than or equal to m and greater than or equal to 0, and in particular the angle C^0 which is the point O. An angle C_λ^r in $\mathfrak{D}$ will be called an ***inner*** angle if one of its points is an inner point of C; otherwise we call it a boundary angle. All angles C_λ^m of the highest dimension in $\mathfrak{D}$ are inner angles.

Let (b_ι, x) be linear forms in R^n, ι running over a finite set of indices I; for every partition of I into three parts K, L, M, consider the angle defined by $(b_\kappa, x)\geqq 0$ $(\kappa\in K)$, $(b_\lambda, x)\leqq 0$ $(\lambda\in L)$, $(b_\mu, x)=0$ $(\mu\in M)$; all those angles, or rather those among them which are different from each other, form a subdivision of R^n. If this process is applied to the set of all linear forms which are needed to define some given angles $C, C', C'', \cdots$ in finite number, then the angles of the resulting subdivision which are contained in C form a subdivision of C; and the same applies to $C', C'', \cdots$.

The intersection of a convex angle of dimension $r\geqq 1$ with the unit-sphere S^{n-1} in R^n, or, as we shall also say, its ***trace*** on S^{n-1}, will be called a ***spherical cell*** of dimension $r-1$. If the angle is degenerate, so is the cell. A nondegenerate cell is homeomorphic to an "element" (a closed simplex) of the same dimension. A degenerate cell is a sphere.

Let Γ be the trace of C^m on S^{n-1}; let $\mathfrak{D}$ be a subdivision of C^m. The traces Γ_λ^{r-1} of the angles C_λ^r of $\mathfrak{D}$ on S^{n-1} for $1\leqq r\leqq m$ form a subdivision of Γ into cells, and so, if $\mathfrak{D}$ is nondegenerate, into topological elements. We can therefore apply elementary results in combinatorial topology to the calculation of the Euler-Poincaré characteristic of such subdivisions.

LEMMA 2. *Let $\mathfrak{D}$ be a nondegenerate subdivision of the angle C^m, consisting*

of the angles C_λ^r $(0 \leqq r \leqq m;\ 1 \leqq \lambda \leqq N_r)$; let N_r' be the number of inner angles of dimension r in $\mathcal{D}$; and write:

$$\chi(\mathcal{D}) = \sum_{r=0}^{m} (-1)^r \cdot N_r, \quad \chi'(\mathcal{D}) = \sum_{r=0}^{m} (-1)^r \cdot N_r'.$$

Then, if C^m is nondegenerate, we have $\chi(\mathcal{D}) = 0$, $\chi'(\mathcal{D}) = (-1)^m$; if C^m is degenerate, $\chi(\mathcal{D}) = \chi'(\mathcal{D}) = (-1)^m$.

This follows at once from the well known value of the characteristic for elements and for spheres, and from the fact that $N_0 = 1$, $N_0' = 0$.

Let now C be a convex angle in R^n, defined as the set of all points $x = \sum_\rho u_\rho \cdot a_\rho$, where the a_ρ are given points and the u_ρ take all values greater than or equal to 0. A linear form (y, x) will be less than or equal to 0 on C if, and only if, $(y, -a_\rho) \geqq 0$ for all ρ; the set of all points y in $\overline{R}^n$ with that property is therefore a convex angle $\overline{C}$. The relationship between C and $\overline{C}$ is easily shown to be reciprocal; we shall say that C and $\overline{C}$ are *dual* to each other. If two angles C, D are such that $C \supset D$, then their dual $\overline{C}$, $\overline{D}$ are such that $\overline{C} \subset \overline{D}$. If an angle is degenerate and reduces to the linear manifold V^r, then its dual is the dual manifold $\overline{V}^{n-r}$. It follows that if $V^r \supset C \supset V^s$, then $\overline{V}^{n-r} \subset \overline{C} \subset \overline{V}^{n-s}$; if, therefore, C is of dimension r and type s, its dual $\overline{C}$ is of dimension $n-s$ and type $n-r$.

LEMMA 3. *Let $\overline{C}$ be the dual of an angle C of type s, and $C \supset V^s$. A point b of $\overline{C}$ is an inner point of $\overline{C}$ if and only if the form (b, x) is less than* 0 *at all points of C other than those of V^s.*

Let C, as above, be the set of points $x = \sum_\rho u_\rho \cdot a_\rho$ when the u's take all values greater than or equal to 0. Then $\overline{C}$ is defined by the inequalities $(y, -a_\rho) \geqq 0$, is of dimension $n-s$, and is contained in the dual $\overline{V}^{n-s}$ to V^s. We have seen that b is an inner point of $\overline{C}$ if and only if $(b, -a_\rho) > 0$ for all those values of ρ for which $(y, -a_\rho)$ does not vanish on $\overline{V}^{n-s}$, that is, for which a_ρ does not lie in V^s; this obviously implies the truth of our lemma.

We now introduce the unit-sphere $\overline{S}^{n-1}$ in $\overline{R}^n$ (to which our earlier remarks about spheres apply); and we shall use the subdivisions of $\overline{S}^{n-1}$, induced by the subdivisions of $\overline{R}^n$ into convex angles, in order to define chains on $\overline{S}^{n-1}$ in the sense of combinatorial topology. All chains should be understood to be $(n-1)$-chains on $\overline{S}^{n-1}$ built up from such subdivisions, the ring of coefficients being the ring of rational integers. We make the usual identifications between certain chains belonging to different subdivisions, by the following rule: if $\mathcal{D}'$ is a refinement of $\mathcal{D}$, and a cell Γ^{n-1} of $\mathcal{D}$ is the union of cells Δ_α^{n-1} of $\mathcal{D}'$, we put $\Gamma^{n-1} = \sum_\alpha \Delta_\alpha^{n-1}$. With that convention, any n-dimensional angle $\overline{C}$ defines a chain, namely, the cell $\Gamma = \overline{C} \cap \overline{S}^{n-1}$, taken with coefficient $+1$ in a suitable subdivision. An angle of dimension less than n is considered as defining the chain 0. Angles being given in $\overline{R}^n$ in finite number, there are

always subdivisions of $\bar{S}^{n-1}$ in which the traces of all those angles appear as chains: we get such a subdivision by making use of all the linear forms which appear in the definition of our angles, as previously explained.

Let C be any convex angle in R^n, and $\bar{C}$ its dual; the chain defined by $\bar{C}$ on $\bar{S}^{n-1}$ will be called the *outer angle* belonging to C, and will be denoted by $\Omega(C)$; that is the chain consisting of the cell $\bar{C}\cap\bar{S}^{n-1}$ if $\bar{C}$ is of dimension n, that is if C is of type 0; if C is of type greater than 0, $\bar{C}$ is of dimension less than n, and $\Omega(C)=0$. With that definition, we have the following theorem:

THEOREM III. *In a subdivision $\mathfrak{D}$ of a convex angle C of dimension m, let C_λ^r $(0\leqq r\leqq m;\ 1\leqq\lambda\leqq N_r')$ be the inner angles; let $\Omega(C)$ and $\Omega(C_\lambda^r)$ be the outer angles belonging to C and to C_λ^r, respectively. Then*:

$$\sum_{r=0}^{m}\sum_{\lambda=1}^{N_r'}(-1)^r\cdot\Omega(C_\lambda^r) = (-1)^m\cdot\Omega(C).$$

We may assume that $\mathfrak{D}$ is nondegenerate, as otherwise C and all C_λ^r are of type greater than 0 and $\Omega(C)=\Omega(C_\lambda^r)=0$. Let Γ be any $(n-1)$-cell in a subdivision of $\bar{S}^{n-1}$ in which $\Omega(C)$ and all $\Omega(C_\lambda^r)$ are sums of cells; put $e=1$ or 0 according as Γ is contained in $\Omega(C)$ or not, and $e_{r,\lambda}=1$ or 0 according as Γ is contained in $\Omega(C_\lambda^r)$ or not. We have to prove that $\sum_{r,\lambda}(-1)^r\cdot e_{r,\lambda}=(-1)^m\cdot e$.

Take first the case $e=1$. Then Γ is contained in the dual $\bar{C}$ of C, and therefore in the duals of all C_λ^r, which all contain $\bar{C}$; all the $e_{r,\lambda}$ are equal to 1, and our formula reduces to $\sum_r(-1)^r\cdot N_r'=(-1)^m$, which is contained in Lemma 2.

Take now the case $e=0$. Let b be an inner point of Γ; call E the angle, or closed half-space, determined by $(b,\ x)\geqq 0$ in R^n; call I the subset of E defined by $(b, x)>0$. As b is not in $\bar{C}$, C has a point in I, and therefore (by Lemma 1) $D=C\cap E$ is an angle of dimension m. Similarly, C_λ^r has a point in I if, and only if, $e_{r,\lambda}=0$, and then $D_\lambda^r=C_\lambda^r\cap E$ is a nondegenerate angle of dimension r. Every inner point of D is an inner point of C, therefore an inner point of a C_λ^r; it must be, then, an inner point of the corresponding D_λ^r, which shows that those D_λ^r which correspond to values of r, λ such that $e_{r,\lambda}=0$ are the inner angles of a subdivision of D; if M_r' is the number of such D_λ^r for a given dimension r, we have therefore, by Lemma 2, $\sum_r(-1)^r\cdot M_r'=(-1)^m$; hence, in that case, $\sum_r\sum_\lambda(-1)^r\cdot e_{r,\lambda}=\sum_r(-1)^r\cdot(N_r'-M_r')=0$, which completes the proof.

Theorem III applies to angles of any dimension and type, and in particular to degenerate angles. Whenever C is of type greater than 0, $\Omega(C)$ is 0.

We observe here that it is merely in order to simplify our exposition that we do not deal with re-entrant, that is, non-convex angles; all our results apply automatically to such angles, provided Theorem III is used to *define* the corresponding outer angles; we mean that, $\mathfrak{D}$ being a subdivision of a non-convex angle C into convex angles, $\Omega(C)$ should be defined by the formula in Theorem III; Theorem III may then be used to show that this $\Omega(C)$ does not

depend upon the choice of $\mathfrak{D}$. Even self-overlapping angles could be treated in the same way.

3. **Dual angles in Euclidean space.** In view of the use to be made of dual angles in §§5–7, we add some remarks on the few circumstances which are peculiar to the case of Euclidean spaces. We therefore assume that a positive-definite quadratic form $\sum_{\mu\nu} g_{\mu\nu}\cdot x^\mu x^\nu$, with constant coefficients $g_{\mu\nu}$, is given in the space R^n of §2. As usual, this is used primarily in order to identify R^n with the dual space $\overline{R}^n$ by means of the formulae $y_\mu = \sum_\nu g_{\mu\nu}\cdot x^\nu$, or, calling $\|g^{\mu\nu}\|$ the inverse matrix to $\|g_{\mu\nu}\|$, $x^\mu = \sum_\nu g^{\mu\nu}\cdot y_\nu$; the two spaces being thus identified, x^μ and y_μ are called the contravariant and the covariant components, respectively, of the vector which they define; they are the same when, and only when, cartesian coordinates are chosen in R^n. We have $(x, x') = \sum_{\mu\nu} g_{\mu\nu}\cdot x^\mu x'^\nu$; two vectors are called orthogonal if $(x, x') = 0$. The unit-sphere $S^{n-1} = \overline{S}^{n-1}$ in R^n is then naturally taken to be the set of all unit-vectors defined by $(x, x) = 1$; only in cartesian coordinates does it appear as $\sum_\mu (x^\mu)^2 = 1$. The dual manifold $\overline{V}^{n-r}$ to a given linear manifold V^r is now the orthogonal or normal manifold to V^r, consisting of all vectors which are orthogonal to every vector in V^r.

Every linear manifold V^r may now itself be regarded as a Euclidean space, and identified with the dual space; if C is an angle in V^r, we may therefore consider its dual *taken within* V^r, which will be an angle in V^r, as well as its dual in R^n. When applied to an angle of given dimension and type, this leads to the following results, which we state in the notation best suited to later applications.

Let R^N be a Euclidean space; let A_r be an angle of dimension n and type r in R^N, contained in the linear manifold T^n and containing the linear manifold T^r; put $q = N - n$, call N^q the orthogonal manifold to T^n, and N^{n-r} the orthogonal manifold to T^r within T^n: the orthogonal manifold to T^r in R^N is then the direct sum $N^{n-r} + N^q$, consisting of all sums of a vector in N^{n-r} and a vector in N^q.

If we take cartesian coordinates w^α $(1 \leqq \alpha \leqq N)$ so that the r first basis-vectors are in T^r, the $n-r$ next ones in N^{n-r}, and the q last ones in N^q, the angle A_r can be defined by $w^{n+\rho} = 0$ $(1 \leqq \rho \leqq q)$ and by a finite number of inequalities of the form $\sum_{\sigma=1}^{n-r} b_\sigma \cdot w^{r+\sigma} \geqq 0$. It is then readily seen that the dual $\overline{A}^{N-r}$ of A_r in R^N, and its dual $\overline{A}^{n-r}$ taken within T^n, are related by the formula: $\overline{A}^{N-r} = \overline{A}^{n-r} + N^q$, which means that $\overline{A}^{N-r}$ consists of all sums of a vector in $\overline{A}^{n-r}$ and a vector in N^q; in other words, a vector is in $\overline{A}^{N-r}$ if and only if its orthogonal projection on T^n belongs to $\overline{A}^{n-r}$. Moreover, $\overline{A}^{n-r}$ is the same as the dual, taken within N^{n-r}, of the trace of A_r on N^{n-r}, that trace being an angle of dimension $n-r$ and of type 0. In this way, questions concerning the dual of an angle of arbitrary dimension and type may be reduced to similar questions concerning the dual of an angle of type 0 and of the highest dimension in a suitable space. The same, of course, could be done in an affine space if desired.

4. **Convex cells and their tubes**(7). We consider an affine space R^N, and its dual $\overline{R}^N$. The linear manifolds which we shall now introduce do not necessarily contain O.

A convex cell in R^N is a compact set of points defined by a finite number of inequalities $(b_\nu, z) \geqq d_\nu$. It is said to be of dimension n if n is the dimension of the smallest linear manifold W^n containing it; it is then known to be homeomorphic to an n-dimensional element. K^n being an n-dimensional cell, contained in the linear manifold W^n, an inner point of K^n is a point, a neighborhood of which in W^n is contained in K^n. Inner points of K^n form an open set in W^n; the closure of that set is K^n, and its complement in K^n, that is, the boundary of K^n, consists of a finite number of convex cells K_λ^r, where r takes all values greater than or equal 0 and less than or equal to $n-1$. We shall count K^n as one of the K_λ^r; with that convention, the K_λ^r, for $0 \leqq r \leqq n$, form a combinatorial complex of dimension n. K_λ^r is a convex cell in a linear manifold W_λ^r; the inner points of K_λ^r are those which belong to no K_μ^s for $s<r$. Every point in K^n is an inner point of one K_λ^r and one only; and, if an inner point of K_λ^r belongs to K_μ^s, then $K_\lambda^r \subset K_\mu^s$.

z being a point in K^n, the points $x = \xi \cdot (z' - z)$, where z' describes K^n and ξ takes all values greater than or equal to 0, form a convex angle, which can be defined by some of the inequalities $(b_\nu, x) \geqq 0$; this will be called the angle of K^n at z; conversely, if x is any point in that angle, $z + \epsilon \cdot x$ will be in K^n for all sufficiently small $\epsilon \geqq 0$. The angle of K^n at z is of dimension n, and contained in the linear manifold V^n, the parallel manifold to W^n through O; if z is an inner point of K_λ^r, the angle of K^n at z is of type r and contains V_λ^r, the parallel manifold to W_λ^r through O; it depends only upon r and λ, and will be denoted by $C_{r,\lambda}$; its dual $\overline{C}_\lambda^{N-r}$ is of dimension $N-r$ and type $N-n$.

Lemma 4. *Let v be a vector in $\overline{R}^N$; v is in $\overline{C}_\lambda^{N-r}$ if, and only if, there is a real number e such that $(v, z) = e$ on K_λ^r and $(v, z) \leqq e$ on K^n; v is an inner point of $\overline{C}_\lambda^{N-r}$ if, and only if, there is an e such that $(v, z) = e$ on K_λ^r and $(v, z) < e$ for all z in K^n except those in K_λ^r.*

As to the first point, let v be in $\overline{C}_\lambda^{N-r}$; let z_0 be in K_λ^r; put $e_0 = (v, z_0)$. For every z in K^n, $z - z_0$ is in $C_{r,\lambda}$, therefore $(v, z - z_0) \leqq 0$, hence $(v, z) \leqq e_0$; therefore e_0 is the least upper bound of (v, z) on K^n and cannot depend upon the choice of z_0 in K_λ^r, so that $(v, z) = e_0$ for all z in K_λ^r; this proves the first point. Conversely, suppose that $(v, z) = e$ for one z in K_λ^r, and that $(v, z') \leqq e$ for all z'

(7) Tubes of convex bodies and of surfaces are of course nothing new, being closely related to the familiar topic of parallel curves and surfaces. On some aspects of this topic which belong to elementary geometry, the reader may consult W. Blaschke, *Vorlesungen über Integralgeometrie*. II, Hamburger Mathematische Einzelschrift, no. 22, Teubner, Leipzig and Berlin, 1937, in particular §37; on p. 93 of that booklet, he will find careful drawings of the tube of a triangle in the plane, and of a tetrahedron in 3-space. The volume of the tube of a closed manifold was recently calculated by H. Weyl, *On the volume of tubes*, Amer. J. Math. vol. 61 (1939) p. 461; part of H. Weyl's calculations will be used in our §6.

in K^n; we have $(v, z'-z) \leqq 0$ for all z' in K^n; this gives $(v, x) \leqq 0$ for all x in $C_{r,\lambda}$, and so v is in $\overline{C}_\lambda^{N-r}$. The second part can now easily be deduced from Lemma 3.

K^n being compact, every linear form (v, z) has on K^n a least upper bound e; the intersection of K^n with the linear manifold $(v, z) = e$ is then one of the cells K_λ^r. This fact, combined with Lemma 4, shows that the angles $\overline{C}_\lambda^{N-r}$ constitute a subdivision of $\overline{R}^N$, according to our definition in §2. The angle C_n of K^n at every inner point is degenerate, and reduces to V^n; its dual $\overline{C}^{N-n}$ is therefore the dual manifold $\overline{V}^{N-n}$ to V^n; the subdivision of $\overline{R}^N$ which consists of the $\overline{C}_\lambda^{N-r}$ is therefore nondegenerate if $N = n$, and degenerate if $N > n$. We leave it as an exercise to the reader to verify that, conversely, every subdivision of R^N into convex angles can be thus derived from a convex cell, or rather from a class of convex cells, in R^N. We observe incidentally that Theorem III of §2 could now be applied; taking $N = n$, which is the only significant case, the $\Omega(\overline{C}_\lambda^{n-r})$ are now the spherical cells determined by the $C_{r,\lambda}$ on the unit-sphere S^{n-1}. In particular, assuming that we are in a Euclidean space, and calling $\mu(C_{r,\lambda})$ the spherical measure of the cell determined by $C_{r,\lambda}$ (which is nothing else than the measure of the "solid angle" $C_{r,\lambda}$), we find that $\sum_{r=0}^{n}\sum_\lambda(-1)^r \cdot \mu(C_{r,\lambda}) = 0$; this is the main result on Euclidean polyhedra in H. Poincaré's paper(8) on polyhedra in spaces of constant curvature; his results on spherical polyhedra could also be derived by similar methods.

Now we take R^N as a Euclidean space, distance and scalar product being defined by means of a fundamental quadratic form (y, y); and we consequently identify $\overline{R}^N$ with R^N, as we did in §3. Let y be any point in R^N; its set-theoretical distance $\delta(y)$ to K^n is a continuous function of y. Let $z = z(y)$ be the nearest point to y in K^n; as K^n is a compact convex set, $z(y)$ is uniquely defined and depends continuously upon y; the vector $v = y - z(y)$, which is of length $\delta(y)$, therefore also depends continuously upon y. That being so, we have $(y-z', y-z') \geqq (v, v)$ for every z' in K^n. Let x be a vector in the angle of K^n at z; $z' = z + \epsilon \cdot x$ is in K^n for sufficiently small $\epsilon \geqq 0$, and then $y - z' = v - \epsilon \cdot x$, so that, for small ϵ, we have $(v-\epsilon \cdot x, v - \epsilon \cdot x) \geqq (v, v)$. That implies that $(v, x) \leqq 0$. If, therefore, z is an inner point of K_λ^r, so that the angle at z is $C_{r,\lambda}$, v is in $\overline{C}_\lambda^{N-r}$. Conversely, let v be in $\overline{C}_\lambda^{N-r}$, and z be an inner point of K_λ^r; as $z' - z$ is in $C_{r,\lambda}$ for every z' in K^n, the same calculation will show that z is the point in K^n nearest to $z + v$.

We now consider the set Θ^N of all points y in R^N whose distance $\delta(y)$ to K^n is at most 1, and we call it the Euclidean tube of K^n in R^N. As Θ^N is a compact convex set and contains an open set in R^N, it is homeomorphic to an N-dimensional closed element. On the other hand, let B^N be the set of all vectors v in R^N such that $(v, v) \leqq 1$, the boundary of which is the unit-sphere S^{N-1}; let $\mathcal{T}(K^n)$ be the subset of the direct product $K^n \times B^N$, consisting of all

(8) H. Poincaré, *Sur la généralisation d'un théorème élémentaire de géométrie*, C. R. Acad. Sci. Paris vol. 140 (1905) p. 113.

elements (z, v) of that product such that, if z is an inner point of K_λ^r, v is in $\overline{C}_\lambda^{N-r}$. We have shown that the relation $y=z+v$ defines a one-to-one bicontinuous correspondence between Θ^N and $\mathcal{T}(K^n)$; the latter, therefore, is a closed subset of $K^n \times B^N$, homeomorphic to B^N; by means of the correspondence defined by $y=z+v$, we identify once for all Θ^N and $\mathcal{T}(K^n)$. Calling $(z(y), v(y))$ the point in $\mathcal{T}(K^n)$ which is thus identified with y in Θ^N, we see that the boundary of Θ^N consists of all points y for which $v(y)$ is on S^{N-1}; in other words, the mapping $y \rightarrow v(y)$ of the tube into B^N maps the boundary into the boundary. As every v is in at least one $\overline{C}_\lambda^{N-r}$, the image of the tube by the mapping $v(y)$ covers the whole of B^N. If we consider a vertex $z_0 = K_\rho^0$ of K^n, and take for v_0 an inner point of the angle $\overline{C}_\rho^N$, all vectors v sufficiently near to v_0 in R^N belong to $\overline{C}_\rho^N$ and to no other angle $\overline{C}_\lambda^{N-r}$, as $\overline{C}_\rho^N$ is an angle of the highest dimension in the subdivision of R^N which consists of the $\overline{C}_\lambda^{N-r}$. Every such vector v, therefore, is the image, by $v(y)$, of the point $y=z_0+v$ and of no other point of Θ^N. This shows that in the neighborhood of such a v_0 the mapping $v(y)$ has the local degree $+1$, and so, as it maps boundary into boundary, it has the degree $+1$ everywhere, provided of course that both Θ^N and B^N are given the orientation induced by that of R^N.

5. **Curved cells and their tubes.** From now onwards, K^n will be a convex cell in an affine space R^n; the object of §§5–6 will be to discuss differential-geometric properties of K^n corresponding to the Riemannian structure determined on it by a certain choice of a ds^2. We write the coordinates in K^n as z^μ $(1 \leqq \mu \leqq n)$; and we choose coordinates ζ^i $(1 \leqq i \leqq r)$ on each one of the cells K_λ^r $(1 \leqq r \leqq n-1)$; for instance, we may choose the ζ^i from among the z^μ, taking care to select such as are independent on K_λ^r, and this may be understood for definiteness, although playing no part in the sequel. In what follows, $N=n+q$ is any integer greater than or equal to n; and we make for §§5–6 the following conventions about the ranges of the various letters which will occur as indices:

$$1 \leqq \alpha \leqq N; \quad 1 \leqq \mu, \nu \leqq n; \quad 1 \leqq i, j \leqq r; \quad 1 \leqq \rho \leqq q; \quad 1 \leqq \sigma \leqq n-r.$$

We shall consider real-valued functions $\phi(z)$, defined on K^n. As usual, such a function is said to be of class C^1 (on K^n) if it has a differential $d\phi = \sum_\mu \phi_\mu(z) \cdot dz^\mu$ with coefficients $\phi_\mu(z) = \partial\phi/\partial z^\mu$ which are continuous functions over K^n; class C^m is defined inductively, ϕ being of class C^m if it is of class C^1 and the $\partial\phi/\partial z^\mu$ are of class C^{m-1}.

Local properties of K^n as a differentiable space are those which remain invariant under a differentiable change of local coordinates with jacobian different from 0. Such properties include the intrinsic definition of the tangent affine space $T^n(z)$ and of the angle of K^n at the point z as follows. $T^n(z)$ is the vector-space consisting of all differentiations $X\phi$, defined over the set of all functions ϕ of class C^1 in a neighborhood of z, which can be expressed as $X\phi = \lim \xi \cdot [\phi(z'') - \phi(z')]$, where z' and z'' both tend to z within K^n, and ξ

tends to $+\infty$. The vectors $X_\mu\phi = \partial\phi/\partial z^\mu$ form a basis for $T^n(z)$, so that every point of $T^n(z)$ can be written as $X\phi = \sum_\mu x^\mu \cdot \partial\phi/\partial z^\mu$; we shall denote by x the point of $T^n(z)$ which, for that basis, has the components x^μ. As in §2, the dual space $\overline{T}^n(z)$ to $T^n(z)$ is the space of the linear forms $(y, x) = \sum_\mu y_\mu \cdot x^\mu$; the elements x of $T^n(z)$ and y of $\overline{T}^n(z)$ are known in tensor-calculus as contravariant and covariant vectors, respectively.

The angle of K^n at z is the subset of $T^n(z)$, consisting of all those differentiations $X\phi$ which can be expressed as $X\phi = \lim \xi \cdot [\phi(z') - \phi(z)]$, where z' tends to z within K^n and ξ tends to $+\infty$; by the correspondence which maps every point $x = (x^\mu)$ in $T^n(z)$ onto the point with coordinates x^μ in the affine space R^n containing K^n, that angle is transformed into the angle of K^n at z as defined in §4, the difference between the two being of course that the latter was defined in affine space whereas the definition of the former refers to K^n as a differentiable space. The relationship between them implies that, if $z = z(\zeta)$ is an inner point of K^r_λ, having in K^r_λ the coordinates ζ^i, the angle at z is of dimension n and type r; we then denote it by $A_{r,\lambda}(\zeta)$; the linear manifold $T^r_\lambda(\zeta)$ contained in $A_{r,\lambda}(\zeta)$ will be identified as usual with the tangent affine space to K^r_λ by the formulae $\partial\phi/\partial\zeta^i = \sum_\mu \partial\phi/\partial z^\mu \cdot \partial z^\mu/\partial\zeta^i$; it is spanned by the r linearly independent vectors $(\partial z^\mu/\partial\zeta^i)$. We denote by $\overline{A}^{n-r}_\lambda(\zeta)$ the dual angle to $A_{r,\lambda}(\zeta)$, which is of dimension $n-r$ and type 0; it is contained in the linear manifold $N^{n-r}_\lambda(\zeta)$ of all vectors $y = (y_\mu)$ such that $(y, x) = 0$ for x in $T^r_\lambda(\zeta)$.

We now consider mappings $f(z) = (f^\alpha(z))$ of K^n into an affine space R^N; $f(z)$ is said to be of class C^m if each $f^\alpha(z)$ is of class C^m. A mapping $f(z) = (f^\alpha(z))$ will be said to define an n-dimensional *curved cell* (K^n, f) if it is of class C^1 and the n vectors $(\partial f^\alpha/\partial z^\mu)$ in R^N are linearly independent for every z in K^n. As usual, the linear manifold spanned by the vectors $(\partial f^\alpha/\partial z^\mu)$ in R^N is identified with the tangent affine space $T^n(z)$ to K^n at z by identifying point $x = (x^\mu)$ in $T^n(z)$ with the vector $(\sum_\mu x^\mu \cdot \partial f^\alpha/\partial z^\mu)$ in R^N; $T^n(z)$ thus appears as imbedded in R^N. The manifold $T^r_\lambda(\zeta)$, as a submanifold of $T^n(z)$ when $z = z(\zeta)$ is in K^r_λ, is thus also imbedded in R^N, and as such is spanned by the vectors $(\partial f^\alpha/\partial\zeta^i) = (\sum_\mu \partial z^\mu/\partial\zeta^i \cdot \partial f^\alpha/\partial z^\mu)$. In the same imbedding, the angle $A_{r,\lambda}(\zeta)$ appears as an angle of dimension n and type r in R^N, contained in $T^n(z)$ and containing $T^r_\lambda(\zeta)$. As the vectors $(\partial f^\alpha/\partial\zeta^i)$ are independent, the mapping f, when restricted to K^r_λ, defines a curved cell (K^r_λ, f) of dimension r in R^N.

We now take R^N as a Euclidean space; cartesian coordinates being chosen for convenience, the distance is defined by the form $(w, w) = \sum_\alpha (w^\alpha)^2$. The quadratic differential form $(df, df) = \sum_\alpha (df^\alpha)^2 = \sum_{\mu,\nu} g_{\mu\nu} \cdot dz^\mu dz^\nu$ is nondegenerate, under the assumptions made on f, and defines a Riemannian geometry on K^n; this amounts to making the tangent affine space $T^n(z)$ into a Euclidean space, either by means of its imbedding in R^N or intrinsically by $(x, x) = \sum_{\mu,\nu} g_{\mu\nu} \cdot x^\mu x^\nu$; the $g_{\mu\nu}$ are functions of z alone. We may then identify $T^n(z)$ with its dual $\overline{T}^n(z)$, as in §3, by the correspondence $y_\mu = \sum_\nu g_{\mu\nu} \cdot x^\nu$; calling, as usual, $\|g^{\mu\nu}\|$ the inverse matrix to $\|g_{\mu\nu}\|$, we have then $x^\mu = \sum_\nu g^{\mu\nu} \cdot y_\nu$; the

y_μ are called the covariant components of the tangent vector x, and the quantities $\sum_\mu x^\mu \cdot \partial f^\alpha/\partial z^\mu$ are its components in R^N.

The Riemannian geometry thus defined in K^n induces on each K^r_λ a Riemannian geometry, with the fundamental form $(df, df) = \sum_{ij} \gamma_{ij} \cdot d\zeta^i d\zeta^j$, where $\gamma_{ij} = \sum_{\mu\nu} g_{\mu\nu} \cdot \partial z^\mu/\partial \zeta^i \cdot \partial z^\nu/\partial \zeta^j$. The determinants of the matrices $\|g_{\mu\nu}\|$, $\|\gamma_{ij}\|$ are denoted by g and γ, respectively; we have $g>0$, $\gamma>0$.

We now call $N^q(z)$ the orthogonal linear manifold to $T^n(z)$ in R^N, that is, the normal linear manifold to the cell at z; and, taking $z=z(\zeta)$ to be an inner point of K^r_λ, we apply to $A_{r,\lambda}(\zeta)$ the results of §3. Identifying, as we now do, $\overline{T}^n(z)$ with $T^n(z)$, the dual linear manifold $N^{n-r}_\lambda(\zeta)$ to $T^r_\lambda(\zeta)$ within $T^n(z)$ appears as the orthogonal manifold to $T^r_\lambda(\zeta)$ within $T^n(z)$, that is, the normal manifold to the subcell (K^r_λ, f); the orthogonal manifold to $T^r_\lambda(\zeta)$ within R^N is then $N^{n-r}_\lambda(\zeta)+N^q(z)$. The dual angle $\overline{A}^{n-r}_\lambda(\zeta)$ to $A_{r,\lambda}(\zeta)$ within $T^n(z)$ is now an angle of dimension $n-r$ and type 0 in the normal manifold $N^{n-r}_\lambda(\zeta)$; it is the same as the dual, taken within $N^{n-r}_\lambda(\zeta)$, of the trace of $A_{r,\lambda}(\zeta)$ on $N^{n-r}_\lambda(\zeta)$. Finally, the dual $\overline{A}^{N-r}_\lambda(\zeta)$ of $A_{r,\lambda}(\zeta)$ within R^N is an angle of dimension $N-r$ and type q, and can be written as $\overline{A}^{N-r}_\lambda(\zeta) = \overline{A}^{n-r}_\lambda(\zeta)+N^q(z)$; this means that a vector w is in $\overline{A}^{N-r}_\lambda(\zeta)$ if, and only if, its orthogonal projection on $T^n(z)$ is in $\overline{A}^{n-r}_\lambda(\zeta)$.

It should be observed that the dual angle $\overline{A}^{n-r}_\lambda(\zeta)$ to $A_{r,\lambda}(\zeta)$, as originally defined in the dual affine space $\overline{T}^n(z)$ to $T^n(z)$, depends only upon K^n regarded as a differentiable space, irrespective of the choice of f or of a Riemannian structure; and we write that a vector y in $\overline{T}^n(z)$, given by its components y_μ, is in $\overline{A}^{n-r}_\lambda(\zeta)$ by writing that $\sum_\mu y_\mu \cdot X(z^\mu) \leqq 0$ for every differentiation X contained in the angle of K^n at $z(\zeta)$. On the other hand, the angles in R^N and in $T^n(z)$ which we have identified with $\overline{A}^{n-r}_\lambda(\zeta)$, and which, for short, we also denote by the same symbol, depend, the former upon the choice of the mapping f, the latter merely upon the $g_{\mu\nu}$.

We now define the *tube* $\mathcal{T}(K^n, f)$ of the curved cell (K^n, f) as the subset of $K^n \times B^N$ which consists of all points (z, w) of that product such that, if z is an inner point of K^r_λ and $z=z(\zeta)$, then w is in $\overline{A}^{N-r}_\lambda(\zeta)$. Whenever f is an affine mapping, that is, when the f^α are linear functions, the tube $\mathcal{T}(K^n, f)$ is the same as the tube $\mathcal{T}(L^n)$ of the convex cell $L^n=f(K^n)$, as defined in §4. Furthermore, if (K^n, f) is an arbitrary cell, the set $\Theta_\delta(K^n, f)$ of all points at a set-theoretical distance δ from $f(K^n)$ in R^N is easily shown to be the same as the set of all points $y^\alpha = f^\alpha(z)+\delta \cdot w^\alpha$ when (z, w) describes $\mathcal{T}(K^n, f)$, and it seems very likely that these relations define a one-to-one correspondence between $\Theta_\delta(K^n, f)$ and $\mathcal{T}(K^n, f)$ provided f itself is a one-to-one mapping and provided δ is sufficiently small.

The central result of this paper is now implicit in the following lemma, which will turn out to contain the Gauss-Bonnet formula for curved cells:

Lemma 5. *The mapping* $(z, w) \to w$ *of the tube* $\mathcal{T}(K^n, f)$ *into* B^N *has everywhere the degree* 1.

The lemma has been proved in §4 for the Euclidean tube of a convex cell. The general case will be reduced to that special case by continuous deformation.

As a preliminary step, we consider the topological space, each point of which consists of a point z in K^n and a set of q mutually orthogonal unit-vectors in $N^q(z)$. This is a fibre-space over K^n, the fibre being homeomorphic to the group of all orthogonal matrices of order q; therefore, by Feldbau's theorem([9]), it is the direct product of K^n with the fibre; that implies that it is possible to choose the q vectors $n_\rho(z)$ as continuous functions of z in K^n so as to satisfy the above conditions for every z. We call $n_\rho^\alpha(z)$ the components of $n_\rho(z)$ in R^N.

Let now $z=z(\zeta)$ be an inner point of K_λ^r, and w a point in R^N; call x, u the orthogonal projections of w on $T^n(z)$ and $N^q(z)$, respectively; call x_μ the covariant components of x, u^ρ the components of u with respect to the basis-vectors $n_\rho(z)$, so that we have

$$w^\alpha = \sum_{\mu,\nu} \frac{\partial f^\alpha}{\partial z^\mu} \cdot g^{\mu\nu} \cdot x_\nu + \sum_\rho n_\rho^\alpha(z) \cdot u^\rho .$$

We have, then, $(w, w)=(x, x)+(u, u)=\sum_{\mu,\nu} g^{\mu\nu}\cdot x_\mu x_\nu+\sum_\rho (u^\rho)^2$; and (z, w) is in the tube $\mathcal{T}(K^n, f)$ if and only if x is in $\overline{A}_\lambda^{n-r}(\zeta)$ and $(w, w)\leqq 1$.

All that applies to the special case when $f^\alpha(z)$ is replaced by $\sum_\mu \delta_\mu^\alpha \cdot z^\mu$, that is, by z^μ for $\alpha=\mu\leqq n$ and by 0 for $\alpha>n$, in which case the tube becomes the Euclidean tube Θ^N of a convex cell; therefore, $z=z(\zeta)$ being again an inner point of K_λ^r, (z, v) will be in Θ^N if and only if the vector in $\overline{T}^n(z)$ with the components $v_\mu=v^\mu$ $(1\leqq\mu\leqq n)$ is in $\overline{A}_\lambda^{n-r}(\zeta)$, and $\sum_\alpha (v^\alpha)^2\leqq 1$. Writing, therefore,

$$x_\mu = \left(\frac{\sum_{\mu,\nu} g^{\mu\nu}(z)\cdot v_\mu v_\nu}{\sum_\mu v_\mu^2}\right)^{-1/2} \cdot v_\mu \qquad (1 \leqq \mu \leqq n),$$

$$u^\rho = v^{n+\rho} \qquad (1 \leqq \rho \leqq q)$$

these formulae, together with the formulae above, define a homeomorphic correspondence between the points (z, v) of the Euclidean tube Θ^N and the points (z, w) of $\mathcal{T}(K^n, f)$.

We now assume coordinates to be such that 0 is in K^n; calling τ a parameter taking the values $0\leqq\tau\leqq 1$, the point $\tau\cdot z=(\tau\cdot z^\mu)$ is in K^n if z is in K^n. For every $\tau>0$, we consider the curved cell $(K^n, \bar{f})$ defined by $\bar{f}(z)=f(\tau\cdot z)/\tau$. Putting $\partial f^\alpha/\partial z^\mu=f_\mu^\alpha(z)$, we have, for the cell $(K^n, \bar{f})$, $\partial \bar{f}^\alpha/\partial z^\mu=f_\mu^\alpha(\tau\cdot z)$, $\bar{g}_{\mu\nu}(z)=g_{\mu\nu}(\tau\cdot z)$, $\bar{g}^{\mu\nu}(z)=g^{\mu\nu}(\tau\cdot z)$, and we may take as normal vectors to that

([9]) J. Feldbau, *Sur la classification des espaces fibrés*, C. R. Acad. Sci. Paris vol. 208 (1939) p. 1621.

cell $\bar{n}_\rho(z) = n_\rho(\tau \cdot z)$. That being so, the above formulae for the transformation of Θ^N into $\mathcal{T}(K^n, \bar{f})$ show that this transformation depends continuously upon τ, and therefore that the tube $\mathcal{T}(K^n, \bar{f})$ is deformed continuously when τ varies. When τ tends to 0, these formulae tend to the corresponding formulae for the cell (K^n, f_0) defined by $f_0^\alpha(z) = \sum_\mu f_\mu^\alpha(0) \cdot z^\mu$ when the normal vectors for (K^n, f_0) are taken as $n_\rho^0(z) = n_\rho(0)$; f_0 being affine, (K^n, f_0) is a convex cell, to which the results of §4 apply.

Lemma 5 follows easily. For the image of our tube in B^N by the mapping $(z, w) \to w$ is deformed continuously when the tube is so deformed; the image of its boundary remains in S^{N-1}. The degree is therefore constant during the deformation; as it is $+1$ for $\tau = 0$, it is $+1$ for $\tau = 1$, which was to be proved.

6. **The Gauss-Bonnet formula for imbedded cells.** We put

$$dw = dw^1 \cdot dw^2 \cdots dw^N.$$

A special consequence of Lemma 5 in §5 is that the integral of dw over the tube $\mathcal{T}(K^n, f)$ is equal to the integral of the same differential form over B^N, that is, to the volume $v(B^N)$ of the interior of the unit-sphere in R^N. Therefore, calling $I_{r,\lambda}$ the integral of $dw/v(B^N)$ over the set of those points (z, w) in the tube for which z is an inner point of K_λ^r, we have

$$\sum_{r=0}^{n} \sum_\lambda I_{r,\lambda} = 1.$$

This becomes the Gauss-Bonnet formula when the $I_{r,\lambda}$ are expressed intrinsically in terms of the Riemannian geometry on K^n. The calculation depends upon a lemma which immediately follows from a formula proved in a recent paper by H. Weyl([10]).

LEMMA 6. *Let* $\|\Lambda_{ij}\|$, $\|L_{ij}^\rho\|$ *be* $q+1$ *matrices of order* r; *and write*

$$\mathrm{P}_{i_1 i_2 j_1 j_2} = \sum_\rho (L_{i_1 j_1}^\rho \cdot L_{i_2 j_2}^\rho - L_{i_1 j_2}^\rho \cdot L_{i_2 j_1}^\rho).$$

Then the integral of $|\Lambda_{ij} + \sum_\rho L_{ij}^\rho \cdot u^\rho| \cdot du^1 \cdot du^2 \cdots du^q$, *taken over the volume* $\sum_\rho (u^\rho)^2 \leq c^2$, *is equal to*:

$$v(B^q) \cdot \sum_{f=0}^{[r/2]} k_{q,f} \cdot c^{q+2f}$$

$$\cdot \sum_{i,j} \frac{\epsilon^{(i)} \epsilon^{(j)}}{2^{2f} \cdot f!(r-2f)!} \mathrm{P}_{i_1 i_2 j_1 j_2} \cdots \mathrm{P}_{i_{2f-1} i_{2f} j_{2f-1} j_{2f}} \Lambda_{i_{2f+1} j_{2f+1}} \cdots \Lambda_{i_r j_r}$$

where $k_{q,f} = 1/(q+2)(q+4) \cdots (q+2f)$, *and the conventions about summation are as explained in* §1.

([10]) Loc. cit., Footnote 7, p. 470. Similar calculations may also be found in W. Killing. *Die nicht-euklidischen Raumformen in analytischer Behandlung*, Teubner, Leipzig, 1885, p. 255.

Our calculation of $I_{r,\lambda}$ will be valid under the assumption that the mapping $f(z)$ is of class C^2; in order, however, to be able to introduce the Riemannian curvature tensor, we assume from now onwards that $f(z)$ is of class C^3. In the course of the calculation of $I_{r,\lambda}$, we simplify notations by omitting the subscript λ.

We may calculate I_r by cutting up the cell K^r into small subsets, and cutting up I_r correspondingly; we take those subsets to be cells of a subdivision of K^r, and so small that it is possible to define, on each of them, q vectors $n_\rho(\zeta)$ of class C^1 and $n-r$ vectors $\nu_\sigma(\zeta)$, also of class C^1, satisfying the following conditions: the $n_\rho(\zeta)$ are an orthonormal basis for the normal linear manifold $N^q(z)$ at $z(\zeta)$; the $\nu_\sigma(\zeta)$ are an orthonormal basis for the normal manifold $N^{n-r}(\zeta)$ to $T^r(\zeta)$ in $T^n(z)$; and, calling ν_σ^α, n_ρ^α the components of those vectors in R^N, the matrix $\Delta = \|\partial f^\alpha/\partial\zeta^i \; \nu_\sigma^\alpha \; n_\rho^\alpha\|$ has a determinant greater than 0. The latter determinant can then be calculated by observing that, if Δ^T is the transpose of Δ and $\Gamma = \|\gamma_{ij}\|$, we have

$$\Delta^T \cdot \Delta = \left\| \begin{matrix} \Gamma & 0 \\ 0 & 1 \end{matrix} \right\|$$

and therefore $(|\Delta|)^2 = \gamma$, so that $|\Delta| = +\gamma^{1/2}$.

$z = z(\zeta)$ belonging to one of our subsets in K^r, let (z, w) be in the tube $T(K^n, f)$. Let x, u be the orthogonal projections of w on $T^n(z)$ and $N^q(z)$, respectively; x is in $\overline{A}^{n-r}(\zeta) \subset N^{n-r}(\zeta)$, so that x can be written as $\sum_\sigma \nu_\sigma(\zeta)\cdot t^\sigma$; let u^ρ be the components of u with respect to the basis n_ρ. We have:

$$w^\alpha = \sum_\sigma \nu_\sigma^\alpha(\zeta)\cdot t^\sigma + \sum_\rho n_\rho^\alpha(\zeta)\cdot u^\rho.$$

As these are functions of ζ, t, u of class C^1, we can express dw in terms of $d\zeta = d\zeta^1\cdot d\zeta^2 \cdots d\zeta^r$, $dt = dt^1\cdot dt^2 \cdots dt^{n-r}$, $du = du^1\cdot du^2 \cdots du^q$:

$$dw = \left| \sum_\sigma \frac{\partial \nu_\sigma^\alpha}{\partial \zeta^i}\cdot t^\sigma + \sum_\rho \frac{\partial n_\rho^\alpha}{\partial \zeta^i}\cdot u^\rho \;\; \nu_\sigma^\alpha \;\; n_\rho^\alpha \right| \cdot d\zeta\cdot dt\cdot du.$$

The determinant is best calculated by multiplying its matrix to the left by Δ^T, the determinant of which has been found to be $+\gamma^{1/2}$; that gives a matrix of the form

$$\left\| \begin{matrix} M & 0 \\ * & 1 \end{matrix} \right\|,$$

which has the determinant $|M|$. That gives:

$$dw = \left| \Lambda_{ij} + \sum_\rho L_{ij}^\rho \cdot u^\rho \right| \cdot \gamma^{-1/2} d\zeta\cdot dt\cdot du$$

if we put

$$\Lambda_{ij} = \sum_{\alpha,\sigma} \frac{\partial f^\alpha}{\partial \zeta^i} \cdot \frac{\partial \nu^\alpha_\sigma}{\partial \zeta^j} \cdot t^\sigma, \qquad L^\rho_{ij} = \sum_\alpha \frac{\partial f^\alpha}{\partial \zeta^i} \cdot \frac{\partial n^\alpha_\rho}{\partial \zeta^j}.$$

In the integration of this, orientation has to be considered. Call t, u the points with the coordinates (t^σ), (u^ρ), respectively, in two auxiliary spaces P^{n-r}, P^q; we also consider the point with the coordinates ζ^i, t^σ, u^ρ in the space $P^N = K^r \times P^{n-r} \times P^q$. The formulae $z = z(\zeta)$, $w = \sum_\sigma \nu_\sigma \cdot t^\sigma + \sum_\rho n_\rho \cdot u^\rho$ define the portion of the tube now under consideration as a homeomorphic image of class C^1 of the subset of P^N defined as follows: ζ is in a given subset of K^r; t is such that $x = \sum_\sigma \nu_\sigma \cdot t^\sigma$ is in $\overline{A}^{n-r}(\zeta)$; and $\sum_\sigma (t^\sigma)^2 + \sum_\rho (u^\rho)^2 \leqq 1$. As $\overline{A}^{n-r}$ depends continuously upon ζ, that set is the closure of an open set in P^N. Call now o_1, o_2, o_3 any orientations of K^r, P^{n-r}, P^q, respectively; the factors in the product $P^N = K^r \times P^{n-r} \times P^q$ being ordered as written, o_1, o_2, o_3 define an orientation $o_1 \times o_2 \times o_3$ in P^N, and therefore a local orientation, also denoted by $o_1 \times o_2 \times o_3$, in the part of the tube which we are discussing. On the other hand, the mappings $(t) \rightarrow \sum_\sigma \nu_\sigma \cdot t^\sigma$, $(u) \rightarrow \sum_\rho n_\rho \cdot u^\rho$ of P^{n-r}, P^q onto $N^{n-r}(\zeta)$, $N^q(z)$ transform o_2, o_3 into orientations, also denoted by o_2, o_3, of $N^{n-r}(\zeta)$, $N^q(z)$. We now choose for o_1, o_2, o_3 the natural orientations of K^r, P^{n-r}, P^q, respectively, defined by the coordinates ζ^i, t^σ, u^ρ taken in each case in their natural order. The condition on the sign of $|\Delta|$ which served to define the ν_σ, n_ρ amounts to saying that the orientations o_1, o_2, o_3 of $f(K^r)$, $N^{n-r}(\zeta)$, $N^q(z)$ at $z = z(\zeta)$ define, when taken in that order, the natural orientation of R^N. That being so, we now show that the local orientation of the tube defined as $o_1 \times o_2 \times o_3$ coincides with that orientation Ω of the tube as a whole which ensures the validity of Lemma 5. That is easily verified for the tube of a convex cell, by identifying it with a subset Θ^N of R^N as in §4. In the general case we use the deformation of our tube into that of a convex cell, by means of which we proved Lemma 5; for, in such a deformation, the manifolds $N^{n-r}(\zeta)$, $N^q(z)$ vary continuously, and therefore we have $o_1 \times o_2 \times o_3 = \Omega$ during the whole deformation, since this is true for one value $\tau = 0$ of the parameter.

We can now proceed to integrate dw by first integrating with respect to u while ζ and t are kept constant; u is to be given all values such that $\sum_\rho (u^\rho)^2 \leqq 1 - \sum_\sigma (t^\sigma)^2$. We first observe that, by differentiating the relations $\sum_\alpha \partial f^\alpha / \partial \zeta^i \cdot \nu^\alpha_\sigma = 0$, $\sum_\alpha \partial f^\alpha / \partial \zeta^i \cdot n^\alpha_\rho = 0$ which express that ν_σ, n_ρ are normal vectors to $T^r(\zeta)$, we get the following expressions for Λ_{ij}, L^ρ_{ij}:

$$\Lambda_{ij} = -\sum_\alpha \frac{\partial^2 f^\alpha}{\partial \zeta^i \partial \zeta^j} \cdot x^\alpha, \qquad L^\rho_{ij} = -\sum_\alpha \frac{\partial^2 f^\alpha}{\partial \zeta^i \partial \zeta^j} \cdot n^\alpha_\rho$$

where $x^\alpha = \sum_\sigma \nu^\alpha_\sigma \cdot t^\sigma$ are the components of the vector x in R^N; these are the negatives of coefficients of the so-called "second fundamental forms" of $f(K^r)$ in R^N. The Λ_{ij} are thus seen not to depend upon the choice of the basis-vectors ν_σ in $N^{n-r}(\zeta)$, but only upon the vector x; as such, we shall now call them $\Lambda_{ij}(x)$; it is known that they are intrinsic quantities with respect to the

Riemannian geometry in K^n, and can be expressed by formula (3) in §1, if we denote by x_μ the covariant components of x; we have $x_\mu = \sum_\alpha \partial f^\alpha / \partial z^\mu \cdot x^\alpha$.

The application of Lemma 6 further leads to the introduction of the quantities

$$P_{i_1 i_2 j_1 j_2} = \sum_\rho (L^\rho_{i_1 j_1} \cdot L^\rho_{i_2 j_2} - L^\rho_{i_1 j_2} \cdot L^\rho_{i_2 j_1}),$$

which also are known to be intrinsic quantities, their expression in terms of the curvature tensor in K^n being given by formula (2) in §1. We now distinguish two cases:

(a) If $r = n$, the ζ^i in the foregoing calculation should be read as z^μ, and γ as g; there are no ν_σ, no t^σ, no Λ_{ij}. Integrating dw first with respect to u, we get, by straightforward application of Lemma 6:

$$I_n = \int_{K^n} \Psi(z) dv(z),$$

where $\Psi(z)$ is defined by formula (1) in §1.

(b) If $r < n$, the integration of dw with respect to u by Lemma 6 gives, if we define the functions $\Phi_{r,f}(\zeta, x)$ by formula (4) in §1

$$v(B^q) \cdot \sum_{f=0}^{[r/2]} k_{q,f} \cdot \left[1 - \sum_\sigma (t^\sigma)^2 \right]^{(q/2)+f} \cdot \Phi_{r,f}(\zeta, x) \cdot dv(\zeta) \cdot dt,$$

where $dv(\zeta) = \gamma^{1/2} \cdot d\zeta$ is the intrinsic volume-element in K^r. We may push the integration one step further, by writing $x = a \cdot \xi$, $x_\mu = a \cdot \xi_\mu$, $t^\sigma = a \cdot \tau^\sigma$, where $\sum_\sigma (\tau^\sigma)^2 = 1$ and $0 \leq a \leq 1$; ξ_μ are thus the covariant components of vector ξ, τ^σ its components with respect to the basis ν_σ, and ξ is on the unit-sphere in $N^{n-r}(\zeta)$; ξ describes a spherical cell $\Gamma(\zeta)$, the trace on that sphere of $\overline{A}^{n-r}(\zeta)$; $\Gamma(\zeta)$ is the outer angle in $N^{n-r}(\zeta)$ of the trace of $A_r(\zeta)$ on $N^{n-r}(\zeta)$. Calling $d\xi$ the area-element or spherical measure on that sphere, we have $dt = a^{r-1} \cdot da \cdot d\xi$. We can now carry out the integration in a, which involves only the elementary integral $\int_0^1 (1-a^2)^{(q/2)+f} \cdot a^{n-2f-1} \cdot da$, and thus find

$$I_r = \int_{K^r} dv(\zeta) \int_{\Gamma(\zeta)} \Psi(\zeta, \xi \mid K^r),$$

where Ψ is defined by formula (5) in §1. This, combined with our earlier result $\sum_{r,\lambda} I_{r,\lambda} = 1$, completes the proof of Theorem II for K^n, with the Riemannian structure defined by the $g_{\mu\nu}$, if we observe that the inner characteristic of K^n is $\chi'(K^n) = (-1)^n$.

It may be observed that, for $r = n-1$, the outer angle $\Gamma(\zeta)$ is reduced to a point, namely, the unit-vector ξ on the outer normal to K^{n-1} in the tangent space to K^n; the integral in $d\xi$ should then be understood to mean the value of the integrand at that point. Similarly, for $r = 0$, K^r is reduced to a point,

and the integral in $dv(\zeta)$ should be understood correspondingly. In the latter case, I_0 contains only one term, corresponding to $f=0$, which is simply the spherical measure of the outer angle $\Gamma(\zeta)$, measured with the area of the sphere taken as the unit. In the case of a Euclidean convex cell, the terms I_0 in our formula are the only ones which do not reduce to 0.

As a preparation to §7, we furthermore have to prove some identities concerning the application of the above results to cells of lower dimension imbedded in K^n. Let L^p be a convex cell, ϕ a one-to-one mapping of class C^3 of L^p into K^n, such that (L^p, ϕ) is a curved cell; we assume that $0 \leqq p \leqq n-1$. For simplicity of notations, we identify L^p with its image in K^n by ϕ, and call (L^p, f) the curved cell which according to earlier conventions should be written as (L^p, h) where h is the product of the two mappings f, ϕ. L^r denoting either L^p or, for $0 \leqq r \leqq p-1$, any one of the boundary cells of L^p, we choose coordinates ζ^i on L^r, and again identify L^r with its image in K^n. The part of the tube of (L^p, f) which corresponds to L^r then consists of all points (z, w) in $K^n \times B^N$ for which $z=z(\zeta)$ is an inner point of L^r, $(w, w) \leqq 1$, and w is in the dual in R^N to the angle of L^p at this point; the latter angle is in the tangent linear manifold to L^p, which as before should be considered as imbedded in the tangent linear manifold $T^n(z)$ to K^n at the same point, and is of dimension p and type r; we denote it by $B_r(\zeta)$. Let $N^{n-r}(\zeta)$ be the normal linear manifold to L^r at ζ; the dual to $B_r(\zeta)$ in R^N consists of all vectors w whose projection x on $T^n(z)$ belongs to the dual to $B_r(\zeta)$ in $T^n(z)$, which is contained in $N^{n-r}(\zeta)$. Let L'^r be an open subset of L^r, so small that we may define on it q vectors n_ρ and $n-r$ vectors ν_σ precisely as before (K^r being replaced by L^r). The calculation and integration of dw for that part of the tube consisting of all points (z, w) with z in L'^r now proceeds, without any change, just as before; the case $r=n$ does not arise, as $r \leqq p \leqq n-1$; calling I_r' the integral of $dw/v(B^N)$ over that part of the tube, we have, therefore:

$$I_r' = \int_{L'^r} dv(\zeta) \int_{\Gamma(L^p,\zeta)} \Psi(\zeta, \xi \mid L^r),$$

where we now denote by $\Gamma(L^p, \zeta)$ the trace on the unit-sphere in $N^{n-r}(\zeta)$ of the dual to $B_r(\zeta)$ in $T^n(z)$. On the other hand, we could have applied our method to L^p itself, considered intrinsically and not as imbedded in K^n; this, for $r=p$, would have given us

$$I_p' = \int_{L'^p} \Psi_0(\zeta) dv(\zeta),$$

if we denote by $\Psi_0(\zeta)$ the invariant built up in L^p just as $\Psi(z)$ was built up in K^n. As this is true for any sufficiently small L'^p, we get, for every inner point ζ of L^p, the identity

$$(6) \qquad \Psi_0(\zeta) = \int_{\Gamma(L^p,\zeta)} \Psi(\zeta, \xi \mid L^p),$$

where $\Gamma(L^p, \zeta)$ being as above defined, is easily seen to be the full sphere in $N^{n-p}(\zeta)$. Similarly, for $0 \leqq r \leqq p-1$, we denote by $\Psi_0(\zeta, \xi_0 \mid L^r)$ the quantity, similar to Ψ, which is built up in L^p from the Riemannian structure defined on L^p by its imbedding in K^n, from the imbedded submanifold L^r, and from a unit-vector ξ_0 normal to L^r in the tangent linear manifold $T^p(\zeta)$ to L^p; and calling $\Gamma_0(L^p, \zeta)$ the trace on the unit-sphere of the dual to $B_r(\zeta)$ in $T^p(\zeta)$, we get as before, ζ being any inner point of L^r:

$$(7) \qquad \int_{\Gamma_0(L^p,\zeta)} \Psi_0(\zeta, \xi_0 \mid L^r) = \int_{\Gamma(L^p,\zeta)} \Psi(\zeta, \xi \mid L^r).$$

The identities (6), (7) contain only quantities which are intrinsic in K^n for the Riemannian structure defined in K^n by the metric tensor $g_{\mu\nu}$. They have just been proved for the case in which the $g_{\mu\nu}$ are defined by a mapping f of K^n into R^N; however, they depend only upon the $g_{\mu\nu}$ and their derivatives of the first and second order at point $z(\zeta)$. It is easy to define a small cell K'^n containing a neighborhood of point $z(\zeta)$ in K^n, and a mapping f' of K'^n into a Euclidean $R^{N'}$, so that (K'^n, f') is a curved cell and that the $g'_{\mu\nu}$ defined by f' over K'^n have, together with their derivatives of first and second order, prescribed values at $z(\zeta)$; in fact, we may do that by taking any analytic $g'_{\mu\nu}$ satisfying the latter conditions, and apply Cartan's theorem[4], but there are of course more elementary methods of obtaining the same result. As (6), (7) are purely local properties of the Riemannian cell K^n and of the imbedded L^p, L^r, they are thus shown to hold without any restriction. They could, of course, be verified by direct calculation; this would be straightforward but cumbersome, and would require another application of Lemma 6.

7. **The Gauss-Bonnet formula for Riemannian polyhedra.** We first define Riemannian polyhedra as follows.

Let P^n be a compact connected topological space, for which there has been given a covering by open subsets Ω_s and a homeomorphic mapping ϕ_s of each Ω_s onto an n-dimensional convex angle C_s which may be R^n; if the ϕ_s and the inverse mappings ψ_s are such that every $\phi_t[\psi_s(x)]$ is of class C^m at every $x \in C_s$ such that $\psi_s(x) \in \Omega_t$, P^n will be called an n-dimensional differentiable polyhedron of class C^m. As noted before (§2), re-entrant angles would lend themselves to similar treatment but are purposely avoided for simplicity's sake.

By a differentiable cell of class C^m, we understand a differentiable polyhedron of class C^m which can be put into a one-to-one correspondence of class C^m with a convex cell.

The beginning of §5 provides a definition for the tangent affine space and the angle of a differentiable cell at any one of its points; those definitions,

being purely local, apply without any change to a differentiable polyhedron. If C is the angle of P^n at the point z, z has a neighborhood homeomorphic to C; if C is of type r, we say that z is of type r in P^n. Points of type n in P^n are called *inner* points of P^n. Points of type at most r (where $0 \leq r \leq n$) form a closed, and therefore compact, subset of P^n, the closure of the set of the points of type r; if the latter consists of N_r connected components, the former is the union of N_r, and not of less than N_r, differentiable polyhedra P^r_λ of dimension r. A point of type r is an inner point of one of the P^r_λ, and of no other P^s_μ; if an inner point of P^r_λ is contained in P^s_μ, then $P^r_\lambda \subset P^s_\mu$. The P^r_λ, for $0 \leq r \leq n-1$, will be called the *boundary polyhedra* of P^n.

By a regular subpolyhedron Q^p in P^n, we understand the one-to-one image of a polyhedron Q^p_0 in P^n, provided it satisfies the following conditions: ζ^i being local coordinates in Q^p_0 at any point, and z^μ local coordinates in P^n at the image of that point, the functions $z^\mu(\zeta)$ which locally define the mapping are of the same class C^m as the polyhedron P^n, and the matrix $\|\partial z^\mu/\partial \zeta^i\|$ is of rank p. Each boundary polyhedron P^r_λ of P^n is a regular subpolyhedron of P^n.

We say that a finite set of distinct regular subpolyhedra Q^r_ρ of P^n forms a *subdivision* $\mathfrak{D}$ of P^n if the following conditions are fulfilled: (a) each point of P^n is an inner point of at least one Q^r_ρ in $\mathfrak{D}$; (b) if Q^r_ρ and Q^s_σ, in $\mathfrak{D}$, are such that there is an inner point of Q^r_ρ contained in Q^s_σ, then $Q^r_\rho \subset Q^s_\sigma$. From (b), it follows that no two polyhedra in $\mathfrak{D}$ can have an inner point in common unless they coincide.

P^n and its boundary polyhedra P^r_λ thus form a subdivision of P^n, which we call the *canonical* subdivision. If $\mathfrak{D}$ is any subdivision of P^n, those polyhedra Q^s_σ in $\mathfrak{D}$ which are contained in a given polyhedron Q^r_ρ in $\mathfrak{D}$ form a subdivision of Q^r_ρ.

LEMMA 7. *If Q^r is a polyhedron in a subdivision $\mathfrak{D}$ of P^n, all inner points of Q^r have the same type in P^n.*

An inner point of Q^r obviously has a type at least r in P^n; hence the lemma is true for $r = n$; we prove it by induction, assuming it to hold for all Q^s_σ in $\mathfrak{D}$ with $s > r$. Let ζ be an inner point of Q^r; call s its type in P^n, so that $s \geq r$; ζ is then inner point of some P^s_λ; we need only show that all points of Q^r, sufficiently near to ζ, are in P^s_λ. That will be the case if all points of P^s_λ, sufficiently near to ζ, are in Q^r; for then, since P^s_λ and Q^r are of class at least C^1 and regular in P^n, we must have $s = r$, and P^s_λ, Q^r must coincide in a neighborhood of point ζ. If that is not so, then ζ must be a limiting point of inner points of P^s_λ which are not in Q^r; as each of the latter points is an inner point of a polyhedron in $\mathfrak{D}$, and there is only a finite number of such polyhedra, it follows that there is a Q^t in $\mathfrak{D}$, such that ζ is a limiting point of inner points of Q^t, each of which is an inner point of P^s_λ and is not in Q^r. This implies that $\zeta \in Q^t$, and therefore $Q^r \subset Q^t$; hence $t > r$, as otherwise an inner point of Q^r

would be inner point of Q^t, and Q^r would be the same as Q^t. By the induction assumption, the lemma holds for Q^t; as there are inner points of Q^t which are inner points of P_λ^s, we have, therefore, $Q^t \subset P_\lambda^s$, and so $Q^r \subset P_\lambda^s$; this proves the lemma.

An immediate consequence is that all the polyhedra, in a subdivision $\mathfrak{D}$ of P^n, which are contained in a given boundary polyhedron P_λ^r of P^n, form a subdivision of that P_λ^r; this can be expressed by saying that every subdivision of P^n is a refinement of the canonical subdivision. In particular, if a polyhedron Q^r, in a subdivision $\mathfrak{D}$ of P^n, contains at least one inner point of P^n, all inner points of Q^r are inner points of P^n; Q^r is then called an *inner polyhedron* of the subdivision.

LEMMA 8. *$\mathfrak{D}$ being a subdivision of P^n, and z any point of P^n, the angles at z of those polyhedra in $\mathfrak{D}$ which contain z form a subdivision of the angle of P^n at z; the inner angles in the latter subdivision are the angles of the inner polyhedra in $\mathfrak{D}$ which contain z.*

In the proof of this lemma, we shall denote by $A(Q)$, Q being any regular subpolyhedron of P^n, the angle of Q at z, if $z \in Q$, and the null-set otherwise. Let x be any vector in $A(P^n)$, defined by an operator $X\phi = \lim \xi \cdot [\phi(z') - \phi(z)]$, where z' tends to z within P^n and ξ tends to $+\infty$; as every z' is an inner point of a Q_ρ^r in $\mathfrak{D}$, and there is only a finite number of such Q_ρ^r, we may define x by a sequence of z', all belonging to one and the same Q_ρ^r; $A(Q_\rho^r)$ then contains x. Let Q^r be a polyhedron of the lowest dimension in $\mathfrak{D}$, such that $x \in A(Q^r)$; if x were not an inner point of $A(Q^r)$, it would be in the angle at z of a boundary polyhedron Q'^s of Q^r, with $s < r$. The polyhedra in $\mathfrak{D}$ which are contained in Q^r form a subdivision of Q^r, and so, by Lemma 7, those which are contained in Q'^s form a subdivision of Q'^s; x would therefore be in the angle at z of one of the latter polyhedra, which would be of dimension at most s, in contradiction with the definition of Q^r. This shows that x is an inner vector of $A(Q^r)$. Suppose, that, at the same time, x is an inner vector of $A(P^n)$; and let x be defined by $X\phi = \lim \xi \cdot [\phi(z') - \phi(z)]$ where the z' are in Q^r; all z', sufficiently near to z, must be inner points of P^n (otherwise x would not be an inner point of $A(P^n)$), and so Q^r must be an inner polyhedron of the subdivision $\mathfrak{D}$. On the other hand, if x is not an inner point of $A(P^n)$, it must be in the angle at z of a boundary polyhedron P_λ^r of P^n; since those polyhedra of $\mathfrak{D}$ which are contained in P_λ^r form a subdivision of P_λ^r, it follows, as above, that x is then an inner point of an angle $A(Q^s)$, where Q^s is a polyhedron in $\mathfrak{D}$ and is contained in P_λ^r.

The proof of the lemma will now be complete if we show that, whenever Q_ρ^r and Q_σ^s belong to $\mathfrak{D}$ and there is an inner point of $A(Q_\rho^r)$ contained in $A(Q_\sigma^s)$, Q_ρ^r itself is contained in Q_σ^s. Using induction, we may, in doing this, assume that the lemma is true for all subdivisions of polyhedra of dimension less than n (the lemma is obviously true when P^n has the dimension 1). The ques-

tion being purely local, we need consider only a small neighborhood of z in P^n, which we may identify with a convex angle in R^n; by the distance of two points in that neighborhood, we understand the Euclidean distance as measured in R^n. Let Q^r be a polyhedron in $\mathfrak{D}$, such that $z \in Q^r$; let x be an inner vector of $A(Q^r)$, defined as above by an operator $X\phi = \lim \xi \cdot [\phi(z') - \phi(z)]$, where we may assume that z' runs over a sequence of inner points of Q^r tending to z. In R^n, the direction of the vector zz' tends to that of the vector x. Our lemma will be proved if, assuming furthermore that x is in the angle at z of a polyhedron in $\mathfrak{D}$ which does not contain Q^r, we show that this implies a contradiction. But the latter assumption implies that, if w is a nearest point to z' in the union W of those polyhedra in $\mathfrak{D}$ which contain z and do not contain Q^r, the direction of the vector zw tends to that of x; we need therefore only show that this implies a contradiction.

w must be contained in a polyhedron Q^s_σ belonging to $\mathfrak{D}$ and containing Q^r, since otherwise it could not be a nearest point to z' in W. Let Q^m_μ be the polyhedron in $\mathfrak{D}$ of which w is an inner point; this is contained in Q^s_σ, and cannot contain Q^r; it is therefore, by Lemma 7, contained in one of the boundary polyhedra Q'^t of Q^s_σ. As there are only a finite number of possibilities for Q^s_σ, Q^m_μ, Q'^t, we may, by replacing the sequence of points z' by a suitable subsequence, assume that these are the same for all w. We now identify a neighborhood of z in Q^s_σ with a convex angle in a Euclidean space R^s; as z, z', w, Q^r, Q'^t are contained in Q^s_σ, we may, in the neighborhood of z, identify them with corresponding points and subsets of that convex angle, and x with the corresponding vector in that same angle.

We have assumed that the direction of the vector zw tends to that of x; therefore Q^r cannot be the same as Q^s_σ, for w is on the boundary of Q^s_σ, and x is an inner vector of Q^r. Therefore Q^r is contained in a boundary polyhedron Q'^u of Q^s_σ; the directions of the vectors zz', zw tend to the direction of x; each point w is in Q'^t, each point z' in Q'^u, and, in the neighborhood of z, Q'^t and Q'^u are the same as two boundary angles of the convex angle Q^s_σ; therefore x must be in the angle at z of $Q'^t \cap Q'^u$, which, by Lemma 7 (applied to Q^s_σ), is the union of polyhedra of $\mathfrak{D}$, so that x is in the angle at z of one of the latter polyhedra. Therefore (applying the induction assumption to Q'^u) Q^r is contained in that polyhedron, and a fortiori in Q'^t. Hence, applying the induction assumption to Q'^t, we get $Q^r \subset Q^m_\mu$, which contradicts an earlier statement.

We now define a *cellular* subdivision of a polyhedron P^n as a subdivision $\mathfrak{D}$, every polyhedron Z^r_ρ in which is a differentiable cell (of the same class as P^n). The application of the results of §6 to arbitrary polyhedra depends upon the following lemma:

LEMMA 9. *Every differentiable polyhedron admits a cellular subdivision.*

This is essentially contained in the work of S. S. Cairns on triangulation,

and also in a subsequent paper of H. Freudenthal on the same subject(11), and need not be proved here.

On a differentiable polyhedron, it is possible to define differentials and differential forms in the usual manner. Such a polyhedron will be called a Riemannian polyhedron if there has been given on it a positive-definite quadratic differential form, locally defined everywhere, in terms of local coordinates z^μ, as $\sum_{\mu,\nu} g_{\mu\nu} dz^\mu dz^\nu$. We make once for all the assumption that our Riemannian polyhedra are of class at least C^3, and that the $g_{\mu\nu}$, which locally define their Riemannian structure, are of class at least C^2 wherever defined. If P^n is such a polyhedron, and Q^p any regular subpolyhedron of P^n, the Riemannian structure of P^n induces again such a structure on Q^p; if ζ^i are local coordinates at a point ζ in Q^p, and the functions $z^\mu(\zeta)$ define the local imbedding of Q^p in P^n at that point, the structure of Q^p at that point is defined by the form $\sum_{i,j}\gamma_{ij}d\zeta^i d\zeta^j$, where $\gamma_{ij}=\sum_{\mu,\nu} g_{\mu\nu}\partial z^\mu/\partial\zeta^i\cdot\partial z^\nu/\partial\zeta^j$. We shall denote by $dv(z)$ the intrinsic volume-element in P^n at z, and by $dv(\zeta)$ the same in Q^p at ζ.

On a Riemannian polyhedron P^n, satisfying the above assumptions, we can define locally at every point z the Riemannian curvature tensor, and hence, by formula (1) of §1, the invariant $\Psi(z)$. Let now Q^p be a regular subpolyhedron of P^n, and ζ a point of Q^p; we shall denote by $N^{n-p}(\zeta)$ the normal linear manifold to Q^p at ζ, which is a submanifold of the tangent space to P^n at ζ. We denote by $\Gamma(Q^p, \zeta)$ the trace, on the unit-sphere, of the dual angle, taken in the tangent space to P^n, of the angle of Q^p at ζ. Furthermore, x being any vector in $N^{n-p}(\zeta)$, we define $\Psi(\zeta, x|Q^p)$ by formulae (2), (3), (4), (5) of §1.

Let now R^s be a polyhedron in a subdivision of Q^p. If $s=p=n$, we define $I(Q^p, R^s)$ as the integral of $\Psi(z)\cdot dv(z)$ over R^s. If $s<n$, we define $I(Q^p, R^s)$ as the integral of $\Psi(\zeta, \xi|R^s)dv(\zeta)$ when ζ describes the set of inner points of R^s and ξ describes, for each ζ, the spherical cell $\Gamma(Q^p, \zeta)$. This implies that $I(Q^p, R^s)=0$ if the inner points of R^s are of type greater than s in Q^p, because $\Gamma(Q^p, \zeta)$ has then a dimension less than $n-s-1$. If, therefore, we consider the sum $\sum_{s,\sigma} I(Q^p, R^s_\sigma)$, taken over all polyhedra R^s_σ of a subdivision of Q^p, this sum has the same value as the similar sum taken for the canonical subdivision of Q^p; the value of that sum is therefore independent of the subdivision by means of which it is defined, and we may write:

$$\sigma(Q^p) = \sum_{s=0}^{p}\sum_{\sigma} I(Q^p, R^s_\sigma),$$

the sum being taken over all polyhedra of any subdivision of Q^p.

(11) See S. S. Cairns' expository paper, *Triangulated manifolds and differentiable manifolds*, in *Lectures in topology*, University of Michigan Conference of 1940, University of Michigan Press, 1941, p. 143, where references will be found to Cairns', Freudenthal's and Whitehead's publications.

$\Psi(z)$ has been defined by using P^n as the underlying Riemannian space. If, on the other hand, we use Q^p as underlying space, we may, substituting p for n in formula (1) of §1 and using the metric and curvature tensors of Q^p, define the similar invariant for Q^p, which we denote by $\Psi_0(\zeta)$. Similarly, ξ_0 being a normal unit-vector to R^s in the tangent space to Q^p at a point ζ of R^s, we define $\Psi_0(\zeta, \xi_0 | R^s)$ by the formulae, similar to (2)–(5) of §1, where Q^p is taken as underlying space instead of P^n. We also define $\Gamma_0(Q^p, \zeta)$ as the trace, on the unit-sphere, of the dual angle, taken in the tangent space to Q^p, of the angle of Q^p at ζ. And we define $I_0(Q^p, R^s)$ as the integral of $\Psi_0(\zeta)dv(\zeta)$ over R^s if $s=p$, and, if $s<p$, as the integral of $\Psi_0(\zeta, \xi_0 | R^s)dv(\zeta)$ when ζ describes the set of inner points of R^s and ξ_0 describes, for each ζ, the spherical cell $\Gamma_0(Q^p, \zeta)$. By the same argument as above, we see, that the sum

$$\sigma_0(Q^p) = \sum_{s,\sigma} I_0(Q^p, R^s_\sigma),$$

taken over all polyhedra R^s_σ of a subdivision of Q^p, is independent of that subdivision. This sum, taken for the canonical subdivision of Q^p, is the same (except for slight changes of notations) as the sum that occurs in the right-hand side of the formula in Theorem II of §1, when that theorem is applied to Q^p. With our present notations, we may, therefore, re-state our Theorem II in the following terms:

THEOREM II. *For every Riemannian polyhedron Q^p, $\sigma_0(Q^p) = (-1)^p \cdot \chi'(Q^p)$.*

We shall first prove that $\sigma(Q^p) = \sigma_0(Q^p)$. As $\sigma(Q^p)$, $\sigma_0(Q^p)$ can be defined from the canonical subdivision of Q^p, it will be enough to prove that, for every Q^r_ρ in that subdivision (that is, either Q^p or one of its boundary polyhedra), $I(Q^p, Q^r_\rho) = I_0(Q^p, Q^r_\rho)$; and this will be proved if we prove that

$$\Psi_0(\zeta) = \int_{\Gamma(Q^p,\zeta)} \Psi(\zeta, \xi | Q^p) \tag{6'}$$

whenever ζ is an inner point of Q^p, and

$$\int_{\Gamma_0(Q^r,\zeta)} \Psi_0(\zeta, \xi_0 | Q^r) = \int_{\Gamma(Q^r,\zeta)} \Psi(\zeta, \xi | Q^r) \tag{7'}$$

whenever ζ is an inner point of a boundary polyhedron Q^r of Q^p. But these identities have been proved, as formulae (6) and (7) of §6, in the particular case when P^n is a Riemannian cell; they are purely local, and depend only upon the angle of Q^p at ζ, the $g_{\mu\nu}$ and their first and second derivatives and the first and second derivatives of the $z^\mu(\zeta)$ at that point; hence they hold in general.

We now prove the important *additivity property* of the function $\sigma(Q^p) = \sigma_0(Q^p)$:

LEMMA 10. *For any subdivision of P^n, the formula holds:*

$$(-1)^n\sigma(P^n) = \sum_{r,\rho}{}'(-1)^r\sigma(Q^r_\rho)$$

where $\sum'$ denotes summation over all inner polyhedra Q^r_ρ of the subdivision.

Call S the sum on the right-hand side. Replacing the $\sigma(Q^r_\rho)$ by their definition, we see that

$$S = \sum_{r,\rho,s,\sigma} \epsilon_{r,\rho,s,\sigma} I(Q^r_\rho, Q^s_\sigma),$$

where the sum is taken over all values of r, ρ, s, σ, and $\epsilon_{r,\rho,s,\sigma}$ has the value $(-1)^r$ whenever $Q^r_\rho \supset Q^s_\sigma$ and Q^r_ρ is an inner polyhedron of the subdivision, and the value 0 otherwise. We may write, therefore:

$$S = \sum_{s,\sigma} J_{s,\sigma},$$

where $J_{s,\sigma}$ is defined by

$$J_{s,\sigma} = \sum_{r,\rho} \epsilon_{r,\rho,s,\sigma} I(Q^r_\rho, Q^s_\sigma);$$

the latter sum may be restricted to those values of r, ρ for which Q^r_ρ contains Q^s_σ and is an inner polyhedron.

We first calculate $J_{s,\sigma}$ in the case $s=n$; the sum then contains only one term, and we have:

$$J_{n,\sigma} = (-1)^n I(Q^n_\sigma, Q^n_\sigma) = (-1)^n I(P^n, Q^n_\sigma).$$

We now take the case $s<n$. From the definition of $I(Q^r_\rho, Q^s_\sigma)$, it follows that $J_{s,\sigma}$ is the integral of $\Psi(\zeta, \xi \mid Q^s_\sigma)dv(\zeta)$ when ζ describes the set of inner points of Q^s_σ, and the integration in ξ, for each ζ, is over the chain:

$$\Delta = \sum_{r,\rho} \epsilon_{r,\rho,s,\sigma} \Gamma(Q^r_\rho, \zeta).$$

Now $\Gamma(Q^r_\rho, \zeta)$, as a chain on S^{n-s-1}, is the same as the outer angle, taken in $N^{n-s}(\zeta)$ according to our definitions in §2, of the trace on $N^{n-s}(\zeta)$ of the angle of Q^r_ρ at ζ. In the sum for Δ, we have all those Q^r_ρ which are inner polyhedra of the subdivision and which contain Q^s_σ, that is, which contain ζ (since ζ is an inner point of Q^s_σ); by Lemma 8, their angles at ζ are the inner angles of a subdivision of the angle of P^n at ζ; since all those angles contain the tangent manifold to Q^s_σ at ζ, their traces on $N^{n-s}(\zeta)$ bear the same relationship to the trace on $N^{n-s}(\zeta)$ of the angle of P^n; we may therefore apply to the outer angles of those traces Theorem III of §2, which gives here $\Delta = (-1)^n\Gamma(P^n, \zeta)$, and therefore:

$$J_{s,\sigma} = (-1)^n I(P^n, Q^s_\sigma),$$

which proves the lemma.

Lemma 10 shows that if Theorem II holds for every cell in a certain cellular subdivision of P^n, it holds for P^n; for, if all Q_ρ are cells and Theorem II holds for them, we have $\sigma(Q_\rho)=(-1)^r\chi'(Q_\rho)=1$, and the right-hand side of the formula in Lemma 10 reduces to the inner characteristic of P^n, as calculated from the given subdivision. Since every polyhedron admits a cellular subdivision, it will now be enough to prove Theorem II for cells. By §6, we know it to hold for an "imbeddable" cell K^n, that is, for one in which the $g_{\mu\nu}$ are defined, as in §6, by a mapping f of K^n into a Euclidean space R^N.

We next take the case of an analytic cell, which we may define by taking a convex cell K^n, and $n(n+1)/2$ functions $g_{\mu\nu}(z)$, analytic over K^n, such that the quadratic form with the coefficients $g_{\mu\nu}(z)$ is positive-definite for every z in K^n. By Cartan's theorem(12), every point of K^n has a neighborhood which can be analytically and isometrically imbedded in a Euclidean space. If, therefore, we subdivide K^n into sufficiently small convex cells (for example, by parallel planes), the Riemannian structure induced on any one of the latter by the given structure in K^n can be defined by an analytic mapping into some Euclidean space, and therefore the results of §6 apply to all cells in that subdivision. Therefore Theorem II holds for K^n.

We now take an arbitrary cell, defined as above by a convex cell K^n and functions $g_{\mu\nu}(z)$ over K^n, the latter being only assumed to be of class C^2; by a theorem of H. Whitney(13), the $g_{\mu\nu}(z)$ can be uniformly approximated, together with their first and second derivatives, by analytic functions and their derivatives. But the expression $\sigma(K^n)$, considered (for a given K^n) in its dependence upon the $g_{\mu\nu}$, depends continuously upon the $g_{\mu\nu}$ and their first and second derivatives; for the integrands Ψ are rational expressions in the $g_{\mu\nu}$, their first and second derivatives, and the components ξ_μ of vector ξ; the denominators in the Ψ consist merely of the determinants g, γ, which are bounded away from 0; $dv(z)$ is $g^{1/2}\cdot dz$, $dv(\zeta)$ is $\gamma^{1/2}\cdot d\zeta$. As to ξ, we may put $\xi_\mu=\omega\cdot\bar{\xi}_\mu$, where $\bar{\xi}_\mu$ describes the trace of the dual of the angle of K^n at ζ on the surface $\sum_\mu(\bar{\xi}_\mu)^2=1$, which is independent of the $g_{\mu\nu}$, and $\omega=(\sum_{\mu\nu}g^{\mu\nu}\cdot\bar{\xi}_\mu\bar{\xi}_\nu)^{-1/2}$. Expressing ξ_μ, $d\xi$ in terms of the $\bar{\xi}_\mu$, we get expressions which are continuous in the $g_{\mu\nu}$. Since $\sigma(K^n)$ is equal to 1 whenever the $g_{\mu\nu}$ are analytic, it follows that it is always 1, and this completes our proof.

Our main result is thus proved in full. Owing, however, to the very unsatisfactory condition in which the theory of differentiable polyhedra has remained until now, the scope of our Theorem II may not be quite adequate for some applications, and we shall add a few remarks which properly belong

---

(12) Loc. cit. Footnote 4.

(13) H. Whitney, *Analytic extensions of differentiable functions defined in closed sets*, Trans. Amer. Math. Soc. vol. 36 (1934) p. 63 (see Lemma 2, p. 69 and Lemma 5, p. 74; as to the latter lemma, which is due to L. Tonelli, cf. C. de la Vallee Poussin, *Cours d'analyse infinitésimale*, vol. 2, 2d edition, Louvain-Paris, 1912, pp. 133–135).

to that theory (to which part of this section may also be regarded as a contribution).

One would feel tempted to regard as a differentiable polyhedron any compact subset of a differentiable manifold which can be defined by a finite number of inequalities $\phi_\nu(z) \geqq a_\nu$, where the ϕ_ν are functions of the same class C^m as the manifold; and one would wish to be able to apply the Gauss-Bonnet formula to such sets.

Now, a compact set P determined, on a manifold M^n of class C^m, by inequalities $\phi_\nu(z) \geqq a_\nu$, the ϕ_ν being functions of class C^m in finite number, actually is a differentiable polyhedron of class C^m, according to our definitions, if the following condition is fulfilled: (A) For any subset S of the set of indices ν, consisting of s elements, and any point z of M^n satisfying $\phi_\sigma(z) = a_\sigma$ for $\sigma \in S$ and $\phi_\nu(z) > a_\nu$ for $\nu \notin S$, the matrix $\|\partial\phi_\sigma/\partial z^\mu\|$ (where σ runs over S and μ ranges from 1 to n) is of rank s. In fact, if condition (A) is fulfilled, let z be any point of P; call S the set of all those indices σ for which $\phi_\sigma(z) = a_\sigma$; by condition (A), their number s is at most n, and we may take the $\phi_\sigma(z)$ as s of the local coordinates at z; the neighborhood of z in P is then an image of class C^m of the angle determined in R^n by the s inequalities $x_\sigma \geqq 0$.

If condition (A) is not satisfied, P need not be a differentiable polyhedron, and indeed it can be shown by examples that "pathological" circumstances may occur. It can be shown, however, that condition (A) is fulfilled, in a suitable sense, for "almost all" values of the a_ν, when the ϕ_ν are given. This gives the possibility of extending the validity of Theorem II to cases when (A) is not fulfilled, by applying it to suitable neighboring values of the a_ν and passing to the limit. Alternatively, almost any "reasonable" definition of a differentiable polyhedron, more general than ours, will be found to be such that our proofs of Lemmas 7 and 8 will remain valid; all our further deductions will then hold provided triangulation is possible.

Finally, it may also be observed that the set P, defined as above by inequalities $\phi_\nu(z) \geqq 0$, can be considered as a limiting case of the set P_ϵ^n defined by the inequalities $\phi_\nu(z) \geqq 0$, $\prod_\nu \phi_\nu(z) \geqq \epsilon$, where ϵ is any number greater than 0. The latter is a polyhedron with a single boundary polyhedron P_ϵ^{n-1} which is a compact manifold of dimension $n-1$; it may be considered as derived from P by "rounding off the edges." We may therefore apply Theorem II to P_ϵ^n; and it is to be expected that the formula thus obtained will tend to a formula of the desired type when ϵ tends to 0. In fact, this idea could probably be used in order to derive our main theorem from the special case of polyhedra P^n bounded by a single $(n-1)$-dimensional manifold.

Haverford College,
Haverford, Pa.

LINEAR OPERATORS IN THE THEORY OF PARTIAL DIFFERENTIAL EQUATIONS

BY

STEFAN BERGMAN

1. Introduction. The taking of the real part of an analytic function of one complex variable is an operation which transforms (in function space) the totality of these functions into the totality of harmonic functions of two variables. Almost every theorem on analytic functions gives rise to a corresponding theorem in the theory of the latter functions. The similarity in structure suggests the use of an analogous approach in the theory of functions satisfying linear partial differential equations of the elliptic type,

$$(1.1)\quad L(U) \equiv U_{z\bar{z}} + a(z, \bar{z})U_z + b(z, \bar{z})U_{\bar{z}} + c(z, \bar{z})U = 0,$$

$$U_{z\bar{z}} = \left(\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2}\right)\Big/4, \qquad U_z = \left(\frac{\partial U}{\partial x} - i\frac{\partial U}{\partial y}\right)\Big/2,$$

$$U_{\bar{z}} = \left(\frac{\partial U}{\partial x} + i\frac{\partial U}{\partial y}\right)\Big/2, \qquad z = x + iy, \qquad \bar{z} = x - iy.$$

In this connection there arises first of all, the question of finding all operators of this kind. *All operators* which transform the class of functions $f(z)$ into the class of functions $U(z, \bar{z})$, $L(U) = 0$, (functions of both classes considered in a sufficiently small neighborhood of the origin) can be determined by formal calculations.

However the transformation of various results requires that the operation be applicable "in the large," that is to say, that every analytic function f regular in a domain $\mathfrak{B}^2$ of a certain class $\mathfrak{D}$ may be transformed into a function U regular in $\mathfrak{B}^2$ and that the inverse operator $U = \mathrm{P}^{-1}(f)$ possess the same property.

Further, for many purposes it is important that for various sequences of functions f_n, the relation $\lim_{n\to\infty} \mathrm{P}[f_n(z)] = \mathrm{P}[\lim_{n\to\infty} f_n(z)]$ shall hold(1).

In addition to the problem of studying all operators and their classification from this point of view, one may consider a particular operator. To a

Presented to the Society, February 22, 1941 under the title *On a class of linear operators applicable to functions of a complex variable* and December 29, 1941 under the title *On operators in the theory of partial differential differential equations and their applications*; received by the editors July 30, 1941, and, in revised form, April 30, 1942.

(1) A system of functions $\{\varphi_\nu(z)\}$, $\nu = 1, 2, \cdots$ possessing the property that every function u regular in a domain $\mathfrak{B}^2$ of certain type $\mathfrak{D}$ can be approximated therein by $\sum_{\nu=1}^{n} \alpha_\nu \varphi_\nu(z)$, may be denoted as a basis of the class of analytic function with respect to $\mathfrak{D}$.

An operator with the above properties transforms a basis $\{\varphi_\nu(z)\}$ into a basis $\{\mathrm{P}[\varphi_\nu(z)]\}$ of the class of functions U with respect to $\mathfrak{D}$. Certain properties of the basis $\{\mathrm{P}(z^\nu)\}$ can be used to characterize the operation P.

certain extent it is useful first to study the latter problem, in order to see how the different properties of the operator influence the transformation of the results, and in order to get a clearer concept of the laws which govern this transformation.

In this paper we shall study the second question. We shall return to the first problem at another place.

NOTATION. We denote the cartesian coordinates of the plane by x, y. Often, however, we shall write $z=x+iy$, $\bar{z}=x-iy$, instead of x and y. We note that if we extend the functions considered to complex values of x and y, the variables z and $\bar{z}$ are no longer conjugate to each other.

Manifolds will be denoted by German letters, the upper index indicating the dimension of the manifold.

$\mathfrak{F}^2$ will always denote a star-domain of the (x, y)-plane, with center at the origin. Its boundary will be denoted by $\mathfrak{f}^1$. $\mathfrak{f}^1$ is supposed to be a differentiable curve.

Further, we denote by $\mathrm{E}[\cdots]$ the set of points whose coordinates satisfy the relations indicated in brackets. $\mathfrak{S}$ means the logical sum.

I. THE OPERATOR AND ITS PROPERTIES

2. The class of functions $\mathfrak{C}(\mathrm{E})$. A survey of obtained results. A complex harmonic function $h(z, \bar{z})$ of two real variables x, y can be represented in the form

$$h(z, \bar{z}) = F(z) + G(\bar{z}) \tag{2.1}$$

where $F(z)$ and $G(z)$ are analytic functions of one complex variable. Since we can write $F(z)=\int_{-1}^{1}f[(z/2)(1-t^2)](1-t^2)^{-1/2}dt$, (2.1) can be written in the form

$$h(z, \bar{z}) = \int_{-1}^{1} \{f[(z/2)(1-t^2)] + g[(\bar{z}/2)(1-t^2)]\}(1-t^2)^{-1/2}dt \tag{2.2}$$

where f and g are analytic functions, of one complex variable, which are regular at the origin.

As was indicated in [3] (numbers in brackets refer to the bibliography), the representation (2.2) can be generalized. Suppose that a, b, c, are analytic functions of two complex variables $z, \bar{z}$. Then for every equation $L(U)=0$ [see (1.1)], there exist functions

$$\mathrm{E}_k(z, \bar{z}, t) = 1 + t^2 z\bar{z}\mathrm{E}_k^*(z, \bar{z}, t), \qquad k = 1, 2, \tag{2.3}$$

such that every solution of $L(U)=0$ can be written in the form

$$\begin{aligned} U(z, \bar{z}) = \int_{-1}^{1}\Big[&\exp\Big(-\int_0^{\bar{z}} a\,d\bar{z}\Big)\mathrm{E}_1(z, \bar{z}, t)f((z/2)(1-t^2)) \\ &+ \exp\Big(-\int_0^{z} b\,dz\Big)\mathrm{E}_2(z, \bar{z}, t)g((\bar{z}/2)(1-t^2))\Big](1-t^2)^{-1/2}dt. \end{aligned} \tag{2.4}$$

For many purposes, instead of considering the functions

$$U(z, \bar{z}) = \exp\left(-\int_0^{\bar{z}} a d\bar{z}\right)\int_{-1}^{1} E_1(z, \bar{z}, t) f((z/2)(1 - t^2))(1 - t^2)^{-1/2} dt,$$

it is useful to investigate the functions

$$(2.5)\quad \begin{aligned} u(z, \bar{z}) &= U(z, \bar{z}) \cdot \left[\exp\left(\int_0^{\bar{z}} a d\bar{z}\right)\right] \\ &= \int E_1(z, \bar{z}, t) f((z/2)(1 - t^2))(1 - t^2)^{-1/2} dt. \end{aligned}$$

Let $\mathfrak{C}(1)$ be the totality of analytic functions of the complex variable z which are regular at the origin. The totality of functions $u(z, \bar{z})$ which can be represented in the neighborhood of the origin in the form

$$(2.6)\quad u(z, \bar{z}) = \mathrm{P}(f) = \int_{-1}^{1} E(z, \bar{z}, t) f((z/2)(1 - t^2))(1 - t^2)^{-1/2} dt, \qquad f \in \mathfrak{C}(1),$$

will be known as the class[2] $\mathfrak{C}(E)$. We define $E(z\ \bar{z}, t)$ to be ***the generating function of*** $\mathfrak{C}(E)$, f ***the associate of*** u, and call the domain in which the representation is valid ***the domain of association.***

If E_1 satisfies a certain partial differential equation, the functions, $U \in [\exp(-\int_0^{\bar{z}} a d\bar{z})] \cdot \mathfrak{C}(E_1)$ satisfy the equation $L(U) = 0$. $[f(z, \bar{z}) \cdot \mathfrak{C}(E_1)$ denotes here the class of functions $f(z, \bar{z}) \cdot u$, where $u \in \mathfrak{C}(E_1)$.$]$ The totality of the solutions of $L(U) = 0$ is given by

$$\left[\exp\left(-\int_0^{\bar{z}} a d\bar{z}\right)\right] \cdot \mathfrak{C}(E_1) + \left[\exp\left(-\int_0^{z} b dz\right)\right] \cdot \mathfrak{C}^*(E_2)$$

where $\mathfrak{C}^*(E)$ is a class analogous to $\mathfrak{C}$, the associates of whose functions are analytic functions of $\bar{z}$. The present paper is devoted to a general study of the functions of any class $\mathfrak{C}(E)$, that is, a class of analytic functions of two real variables x, y which can be represented in a sufficiently small neighborhood of the origin by the right-hand member of (2.6)[3].

In this paper we drop the assumption that the functions $u \in \mathfrak{C}(E)$ satisfy

---

(2) We may also consider classes of functions for which a representation analogous to (2.6) holds in the neighborhood of a point a, $a \neq 0$. Functions satisfying $L(U) = 0$ possess the property that the representation (2.4) exists for every point a. The study of the dependence upon the point a of $E(z, \bar{z}, t | a)$ and the associate $f(z | a)$, of a function u is an interesting problem of the theory.

(3) The relation (2.6) may be interpreted as a mapping (in the function space) of $\mathfrak{C}(1)$ into the class $\mathfrak{C}(E)$. We are going to study the duality between the theories of the functions of $\mathfrak{C}(1)$, and those of $\mathfrak{C}(E)$.

Note that our space of functions includes those which are not defined *in one fixed domain, but only in a sufficiently small neighborhood of the origin.*

These functions arise also in other connections, for example, as a set of particular solutions

a linear partial differential equation. We suppose only that the functions u possess the two properties A and B which we now describe.

A. The function E can be written in the form

$$\text{(2.7)} \qquad E(z, \bar{z}, t,) = 1 + t^2 z\bar{z}E^*(z, \bar{z}, t),$$

where E^* is an analytic function of two *complex* variables z, $\bar{z}$ regular in the region $E[|z| < \infty, |\bar{z}| < \infty]$ and a continuously differentiable function of z, $\bar{z}$, t in $E[|z| < \infty, |\bar{z}| < \infty, |t| \leq 1]$.

We note that from A follow:

A_1. Every $u \in \mathcal{C}(E)$, regular in a star-domain $\mathfrak{F}^2$, can be continued analytically in(4)

$$\text{(2.8)} \qquad \mathfrak{R}^4(\mathfrak{F}^2) = E[z = a + ib, \bar{z} = a - ib, (a, b) \in \mathfrak{F}^2, a, b \text{ real}].$$

A_2. For every $u \in \mathcal{C}(E)$,

$$\text{(2.9)} \qquad |u(z, \bar{z})| \leq c \max_{\zeta \in \mathfrak{f}^1} |f(\zeta)|, \qquad \text{for } (z, \bar{z}) \in \mathfrak{R}^4(\mathfrak{F}^2)$$

and

$$c = \max_{(z, \bar{z}) \in \mathfrak{R}^4(\mathfrak{F}^2), |t| \leq 1} |E(z, \bar{z}, t)|.$$

B. There exists, for every $\mathfrak{F}^2$, an operator $G(z, \bar{z}, \zeta, \bar{\zeta}, X_{00}, \cdots, X_{mn})$ such that

$$\text{(2.10)} \quad u(z, \bar{z}) = \int_{\mathfrak{f}^1} G[z, \bar{z}, \zeta, \bar{\zeta}, u(\zeta, \bar{\zeta}), u_\zeta(\zeta, \bar{\zeta}), \cdots, u_{\zeta^m \bar{\zeta}^n}(\zeta, \bar{\zeta})] ds_\zeta, \ (z, \bar{z}) \in \mathfrak{R}^4(\mathfrak{F}^2).$$

Here $\mathfrak{f}^1$ is the boundary of $\mathfrak{F}^2$, ds_ζ is the line element of $\mathfrak{f}^1$, and $G(z, \bar{z}, \zeta, \bar{\zeta}, X_{00} \cdots X_{mn})$, $[|X_{pq}| < \infty, (pq) = (00), \cdots, (mn)]$ is an analytic function of two complex variables z, $\bar{z}$, which is regular in $\mathfrak{R}^4(\mathfrak{F}^2)$.

Since we suppose that $E(z, \bar{z}, t)$ is an analytic function of two complex variables z, $\bar{z}$ the functions $u(z, \bar{z})$ are also analytic functions of two complex variables. In general, they can be continued analytically in the space ranged over by two complex variables and, therefore, outside of their domains of association.

In developing the theory of the functions of $\mathcal{C}(E)$, one may distinguish the following two types of results:

---

of partial differential equations of order higher than two, or as solutions of systems of partial differential equations. We note that often the pair of solutions of a system of equations may be interpreted physically, for example, as the stream and potential function of a flow.

It should be stressed that our investigations concern the behavior of functions $u(z, \bar{z})$ for real values of x and y (that is, for z and $\bar{z}$ which are conjugate). However, in some auxiliary considerations we shall extend x and y to complex values.

(4) To every point with the coordinates $x = a$, $y = b$ there correspond planes $\mathfrak{P}^2(a, b) = E[z = a + ib]$ and $\mathfrak{Q}^2(a, b) = E[\bar{z} = a - ib]$ in the four-dimensional space. Thus, $\mathfrak{R}^4(\mathfrak{F}^2)$ is the intersection of two four-dimensional cylinders.

(1) Theorems in which $u(z, \bar{z})$ is considered inside of the domain of association, $\mathfrak{A}^2$.

(2) Theorems concerning the behavior of u on the boundary(5) of $\mathfrak{A}^2$, as well as the properties of u outside of $\mathfrak{A}^2$.

Many theorems of the type (1) follow immediately, for the functions u, from corresponding results in the theory of analytic functions, by using the representation(6) (2.6) and the Corollary 3.1 (p. 136). In particular this is true for many theorems stating that an analytic function can be represented as the sum of a linear combination of a finite or infinite number of analytic functions, belonging to a given set. For instance, this is true for theorems dealing with development in series, and on approximation, and the Cauchy integral formula, as well as the many consequences of these theorems. (See §§5 and 6.)

In §7 we show that the connection between the position of certain singularities of $u(z, \bar{z})$ and the coefficients B_{mn} of the development $\sum B_{mn}x^m y^n$ of u is, to a certain degree independent of any special choice of E^* (see (2.7), [2] and [3]). The same holds for various theorems concerning the connection between B_{mn} and the regularity domain, the growth of u and averages of u with certain weight functions. In §9 we study the coincidence, along curves, of the values of functions belonging to two different classes. These considerations show that many properties of the functions u of the class $\mathcal{C}(\mathrm{E})$ are either independent of the choice of E^* (see (2.7)) or depend upon E^* in a simple manner.

In particular since functions u satisfying (1.1) can be presented in the form (2.4) with E_k of the form (2.7) (see [3, §1])(7) these results are valid for the solution of partial differential equations

$$L(U) = 0.$$

Since E_k^* is the only expression in (2.3) which depends on a, b, c, *the resulting relations are independent of the coefficients a, b, c of the equation.*

On the other hand, solutions of certain equations $L(U)=0$ form also a class $\mathcal{C}(\mathrm{E})$ wherein E is of a quite different form from that here considered; for instance wherein

(5) The study of singularities of functions of $\mathcal{C}(\mathrm{E})$ is a particular one of this group of questions.

(6) In previous papers [3], [4] we proved that the solutions of an equation $L(u)=0$ can be presented in the form (2.4) with E_k, $k=1, 2$ possessing properties A and B. With this result we constructed two sets of functions $\{\phi_\nu^{(k)}(z, \bar{z})\}$ which serve as bases of the class $\mathcal{C}(\mathrm{E}_k)$ with respect to the star-domains. In [4] we discussed the application of this method to the actual solution of boundary value and characteristic value problems.

(7) The existence of the operator $G(z, \bar{z}, \zeta, \bar{\zeta}, X_{00}, X_{10}, X_{01})$ for these functions follows from Green's formula. (See [11, p. 515, (9)].)

We note that U differs slightly from u. (See (2.4) and (2.5).) This should be kept in mind when formulating results for solutions of differential equations.

$$\mathbf{E} = \left[\sum_{\nu=1}^{m} t^\nu q_\nu(z, \bar{z}) \right] \left\{ \exp \left[\sum_{\nu=1}^{n} t^\nu p_\nu(z, \bar{z}) \right] \right\}. \tag{2.11}$$

(See [3, §3].)

Since various properties of the class $\mathcal{C}(\mathbf{E})$ depend to a large extent upon $\mathbf{E}$, the study of $\mathcal{C}(\mathbf{E})$ with various forms of $\mathbf{E}$ gives results which are quite different from the results of this paper[8].

The study of $\mathcal{C}(\mathbf{E})$ with $\mathbf{E}$ of the form (2.11) seems to be particularly important for the study of singularities of functions satisfying $L(u)=0$.

It is possible to show that to a pole of the associate function f (of u) there corresponds in this case, a singularity of u with the following property: u satisfies two ordinary differential equations in z and $\bar{z}$, with coefficients which depend in a simple manner upon p_n and q_n. (See [3, §2] and [2].)

3. **Determination of f in terms of $u(z, \bar{z})$, $(z, \bar{z}) \in \mathfrak{R}^4(\mathfrak{F}^2)$.** In this section we shall determine the operator

$$f(\zeta) = R(\zeta \mid u), \tag{3.1}$$

inverse to (2.6). For the sake of simplicity, we shall deal in the future with a certain operator Q instead of R. Q is connected with R by the relation[9]

$$R(\xi \mid u) = \frac{(2\xi)^{1/2}}{\Gamma(1/2)} \frac{d^{1/2} Q(2\xi \mid u)}{d\xi^{1/2}}, \tag{3.2}$$

which may be written in the form of an integral relation:

$$R(\xi \mid u) = \frac{2}{\pi} \int_0^{\pi/2} \xi \sin \vartheta \, \frac{dQ(2\xi \sin^2 \vartheta \mid u)}{d(\xi \sin^2 \vartheta)} \, d\vartheta + \pi^{-1} u(0, 0) \tag{3.3}$$

(cf. [3, p. 1177]).

Theorem 3.1. *We have the relation*

$$Q(\xi \mid u) = u(\xi, 0). \tag{3.4}$$

Proof. It follows from property A (cf. p. 133), that

---

[8] As shown in [3], if $\mathbf{E}(z, \bar{z}, t)$ satisfies a certain partial differential equation, then U, $U \in \mathcal{C}(\mathbf{E})$, satisfies the equation $L(U)=0$. Clearly this equation may have many solutions, any of which can be used as $\mathbf{E}$. We note that the equation (1.2) of [3] can be simplified. Introducing $p = tz^{1/2}$ instead of t, equation (1.2) becomes $\mathbf{E}^*_{\bar{z}p} - p^{-1}\mathbf{E}^*_{\bar{z}} + 2p(\mathbf{E}^*_{z\bar{z}} + D\mathbf{E}^*_z + F\mathbf{E}) = 0$. $\mathbf{E}^*$ is connected with $\mathbf{E}$ by the relation (1.8) of [3]. (See also [3, p. 1177].) Finally, writing $\mathbf{E}^* = 1 + pQ(z, \bar{z}, p)$ we obtain $Q_{\bar{z}p} + 2p(Q_{z\bar{z}} + DQ_{\bar{z}} + FQ) + 2F = 0$.

For certain purposes it is also useful to consider an operator of the form $\mathbf{P}(f) = f(z)\mathbf{E}_1(z, \bar{z}) + \int_\alpha^\beta \mathbf{E}_2(z, \bar{z}, s) f(p(z, s)) ds$ or, others of more complicated structure.

Note that the above operator transforms $\log z$ into a function with a logarithmic singularity.

[9] We define, as usual, $d^{1/2}(\sum \alpha_n \zeta^n)/d\zeta^{1/2} = \sum (\Gamma(n+1)/\Gamma(n+1/2)) \alpha_n \zeta^{n-1/2}$.

$$u(z, 0) = \int_{-1}^{1} f((z/2)(1-t^2))(1-t^2)^{-1/2} dt. \tag{3.5}$$

Suppose now that $f(z) = \sum_{n=0}^{\infty} a_n z^n$. Then

$$\begin{aligned} u(z, 0) &= \sum_{n=0}^{\infty} a_n (z/2)^n \int_{-1}^{1} (1-t^2)^{n-1/2} dt \\ &= \sum_{n=0}^{\infty} a_n (z/2)^n \Gamma(n+1/2)\Gamma(1/2)/\Gamma(n+1). \end{aligned} \tag{3.6}$$

Since $\boldsymbol{R}(z \mid u) = f(z) = \sum_{n=0}^{\infty} a_n z^n$, the relation (3.4) follows from (3.2) and (3.6).

If $f(z/2)$ is regular in $\mathfrak{F}^2$, then it follows, by (2.6) and A, that $u(z, \bar z)$, too, is regular in $\mathfrak{F}^2$. Theorem 3.1 yields the inverse statement given by

COROLLARY 3.1. *Suppose that $u(z, \bar z)$, $u \in \mathcal{C}(\mathrm{E})$, is regular in a star-domain $\mathfrak{F}^2$. Then $f(z/2)$ is regular in $\mathfrak{F}^2$.*

This fact is an immediate consequence of (3.4). For the regularity of $f(z/2) = \boldsymbol{R}(z/2 \mid u)$ in $\mathfrak{F}^2$ follows by (3.2) from the regularity of $\boldsymbol{Q}(z \mid u)$ in the same domain. The regularity of $\boldsymbol{Q} \equiv u(z, 0)$ in $\mathfrak{F}^2$ follows from B since the domain $\mathrm{E}[z \in \mathfrak{F}^2, \bar z = 0]$ lies in $\mathfrak{R}^4(\mathfrak{F}^2)$.

4. Determination of the associate function in terms of $u(z, \bar z)$ in the real plane. Relations (3.3) and (3.4) give the representation of the associate function. But in this formula there appear functions $u(z, \bar z)$ for which the values of z and $\bar z$ are, in general, not conjugate to each other. This means that we consider $v(x, y) = u(z, \bar z)$ for complex values of x, y. On the other hand, for many questions it is important to have a formula where $v(x, y)$ appears and takes on only real values of the arguments. We obtain such a formula by substituting the right-hand member of (2.10) for u in (3.4). (See [5, §3].) However, this last formula is inconvenient because the expression obtained for $\boldsymbol{Q}$ depends on G, and therefore on $\mathcal{C}(\mathrm{E})$. Because of the importance of formulas for $\boldsymbol{R}$, we shall indicate other expressions for $\boldsymbol{Q}$ which are independent of $\mathcal{C}(\mathrm{E})$.

THEOREM 4.1. *Suppose that $u(z, \bar z) \in \mathcal{C}(\mathrm{E})$ is regular in $\mathrm{E}[x^2+y^2 \leq 4R^2]$. Assume $r < R$. Then*

$$\boldsymbol{Q}(\zeta \mid u) = \frac{1}{2\pi}\int_0^{2\pi} \left\{ \sum_{k=0}^{\infty} (-1)^k \frac{r^k}{k!} \frac{\partial^k}{\partial r^k}\left[\frac{u(re^{i\varphi}, re^{-i\varphi})re^{i\varphi}}{re^{i\varphi} - \zeta}\right]\right\} d\varphi. \tag{4.1}$$

Proof. By (2.7) we have

$$u(z, \bar z) = \boldsymbol{Q}(z \mid u) + \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} A_{mn} z^m \bar z^n, \qquad \boldsymbol{Q}(z \mid u) = \sum_{m=0}^{\infty} A_{m0} z^m. \tag{4.2}$$

The series (4.2) converges uniformly and absolutely(10) in $\mathrm{E}[|z| < R, |\bar z| < R]$.

(10) The absolute convergence follows from $\sum_{m=0}^{\infty}\sum_{n=1}^{\infty} |A_{mn}| |z^m \bar z^n| \leq \int_{-1}^{1} [\sum_{m=0}^{\infty}\sum_{n=1}^{\infty} |\mathrm{E}_{mn}(t)| \cdot |z^m \bar z^n|][\sum_{n=0}^{\infty} |a_n z^n| ((1-t^2)/2)^n](1-t^2)^{-1/2} dt$.

Integrating along the circle $|z| = r$ of the real plane, we obtain

$$(4.3)\qquad \frac{1}{2\pi}\int_0^{2\pi}\left\{\frac{\partial^k}{\partial r^k}\left[\frac{u(re^{i\varphi}, re^{-i\varphi})}{r^q}\right]\right\}\frac{d\varphi}{e^{iq\varphi}} = \frac{\partial^k}{\partial r^k}\left[\sum_{\nu=0}^{\infty} A_{\nu+q,\nu} r^{2\nu}\right].$$

Now $\sum_{\nu=0}^{\infty} A_{\nu+q,\nu}\zeta^{2\nu}$ is an analytic function of the complex variable ζ, $|\zeta| \leq R$. We introduce a new variable $Z = \zeta - r$, and develop this function about the point $Z = 0$. Then putting $Z = -r$, we obtain

$$(4.4)\qquad A_{q0} = \frac{1}{2\pi}\int_0^{2\pi}\left\{\sum_{k=0}^{\infty}(-1)^k\frac{r^k}{k!}\frac{\partial^k}{\partial r^k}\left[\frac{u(re^{i\varphi}, re^{-i\varphi})}{r^q}\right]\right\}\frac{d\varphi}{e^{iq\varphi}}.$$

Equation (4.1) follows from (4.4).

REMARK. Analogous considerations yield $Q(z|u)$ in terms of $\mathrm{Re}(u)$ and $\mathrm{Im}(u)$, respectively. In fact, by (4.2) we have

$$u(z, \bar{z}) + \bar{u}(\bar{z}, z) = Q(z \mid u) + \overline{Q(z \mid u)} + \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} A_{mn} z^m \bar{z}^n + \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \bar{A}_{mn}\bar{z}^m z^n,$$

and

$$(1/2\pi)\int_0^{2\pi}(u + \bar{u})d\varphi = (A_{00} + \bar{A}_{00}) + (A_{11} + \bar{A}_{11})r^2 + (A_{22} + \bar{A}_{22})r^4 + \cdots,$$

$$(1/2\pi)\int_0^{2\pi}(u + \bar{u})e^{-iq\varphi}d\varphi = A_{q0}r^q + (A_{q+1,1} + \bar{A}_{1,q+1})r^{q+2} + (A_{q+2,2} + \bar{A}_{2,q+2})r^{q+4} + \cdots$$

for $q \geq 1$. Similarly,

$$(1/2\pi)\int_0^{2\pi}(u - \bar{u})d\varphi = (A_{00} - \bar{A}_{00}) + (A_{11} - \bar{A}_{11})r^2 + \cdots.$$

In the same way as before, the associate function

$$(4.5)\qquad f(z) = T(z \mid u^{(1)}) + iC(z \mid u^{(1)}) = G(z \mid u^{(2)}) + iM(z \mid u^{(2)})$$

can be determined (to within A_{00}) from either the real or imaginary part of $u = u^{(1)} + iu^{(2)}$.

For some purposes, it is convenient to have a formula for $Q(z|u)$ in which no derivatives of u appear. In order to obtain such an expression, we need certain lemmas.

LEMMA 4.1. *Let* $\mathfrak{r}^1 = E[0 \leqq r_1 \leqq r \leqq r_2 < R]$. *There exists a set of functions* $\phi_\nu(z)$ *such that*

$$\iint_{|z|<R} \phi_\nu(z)\phi_\mu(z)dxdy = 1, \quad \textit{for} \quad \nu = \mu, \tag{4.6a}$$
$$= 0, \quad \textit{for} \quad \nu \neq \mu,$$

$$\int_{\mathfrak{r}^1} \phi_\nu(r)\bar{\phi}_\mu(r)dr = \lambda_\nu \quad \textit{for} \quad \nu = \mu, \tag{4.6b}$$
$$= 0 \quad \textit{for} \quad \nu \neq \mu.$$

Proof. As is well known, the system $\{(n/\pi)^{1/2}Z^{n-1}\}$ is orthonormal in the unit circle, $|Z| < 1$. Set $h(Z) = \sum a_n(n/\pi)^{1/2}Z^{n-1}$, $g(Z) = \sum b_n(n/\pi)^{1/2}Z^{n-1}$. The Hermitian form

$$\begin{aligned}\int_{\rho_1}^{\rho_2} h(\rho)\bar{g}(\rho)d\rho &= \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_m\bar{b}_n \frac{(nm)^{1/2}}{\pi}\int_{\rho_1}^{\rho_2}\rho^{n+m-2}d\rho \\ &= \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_m\bar{b}_n \frac{(nm)^{1/2}}{\pi(n+m-1)}(\rho_2^{n+m-1} - \rho_1^{n+m-1}), \rho_1 < \rho_2 < 1,\end{aligned} \tag{4.7}$$

is completely continuous (vollstetig), since $\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\rho_2^{n+m-1}$ exists (see [9, pp. 147–151, especially p. 151]). Therefore, (4.7) can be written in the form

$$\sum_{\nu=1}^{\infty} \lambda_\nu^*\left(\sum_{m=1}^{\infty} \mathfrak{o}_{\nu m}a_m\right)\overline{\left(\sum_{n=1}^{\infty} \mathfrak{o}_{\nu n}b_n\right)} \tag{4.8}$$

where $\{\mathfrak{o}_{\nu n}\}$ is a unitary matrix. Since $\{\mathfrak{o}_{\nu n}\}$ is unitary, the functions

$$\psi_\nu(Z) = \sum_{n=1}^{\infty}(n/\pi)^{1/2}\mathfrak{o}_{\nu n}Z^{n-1} \tag{4.9}$$

have the property that

$$\iint_{|Z|<1}\psi_\nu(Z)\overline{\psi_\mu(Z)}dXdY = 1, \quad \int_{\rho_1}^{\rho_2}\psi_\nu(\rho)\bar{\psi}_\mu(\rho)d\rho = \lambda_\nu^* \text{ for } \nu = \mu, \tag{4.10}$$
$$= 0, \qquad\qquad = 0 \text{ for } \nu \neq \mu.$$

Substituting in (4.9) $Z = z/R$, $\rho_k = r_k/R$, $\lambda_\nu = \lambda_\nu^*/R$, $\phi_\nu(z) = \psi_\nu(z/R)/R$, we obtain a system of the desired form.

LEMMA 4.2. *Let* $g(z)$ *be an analytic function of one complex variable* z, *regular in* $|z| \leqq R$, *which takes on the values* $F(r)$ *on* $\mathfrak{r}^1$, *and for which* $\iint_{|z|<R}|g(z)|^2dxdy < \infty$. *Then*

$$g(z) = \sum_{\nu=1}^{\infty}\frac{\phi_\nu(z)}{\lambda_\nu}\int_{\mathfrak{r}^1}F(r)\bar{\phi}_\nu(r)dr, \qquad |z| < R. \tag{4.11}$$

Proof. By theorems on orthogonal functions [7 p. 26], it is seen that $g(z)$ can be represented in $|z|<R$ in the form $g(z)=\sum_{\nu=1}^{\infty}A_\nu\phi_\nu(z)$, where the series converges uniformly in $|z|\leqq\rho<R$. Since $\sum_{\nu=1}^{\infty}A_\nu\phi_\nu(r)=F(r)$ for $r\in\mathfrak{r}^1$, the relation

$$\int_{\mathfrak{r}^1}\bar{\phi}_\mu(r)\sum_{\nu=1}^{\infty}A_\nu\phi_\nu(r)dr=\sum_{\nu=1}^{\infty}A_\nu\int_{\mathfrak{r}^1}\phi_\nu(r)\bar{\phi}_\mu(r)dr=\int_{\mathfrak{r}^1}F(r)\bar{\phi}_\mu(r)dr$$

gives us $\lambda_\mu A_\mu=\int_{\mathfrak{r}^1}F(r)\bar{\phi}_\nu(r)dr$, which yields (4.11).

REMARK. From (4.11) and (4.9) we have

$$g(0)=R^{-1}\sum_{\nu=1}^{\infty}(\mathfrak{o}_{\nu 1}/\lambda_\nu)\int_{\mathfrak{r}^1}F(r)\bar{\phi}_\nu(r)dr. \tag{4.12}$$

THEOREM 4.2. *Under hypothesis of Theorem* 4.1, *we have*

$$Q(\zeta\mid u)=\frac{1}{2\pi R}\left[\sum_{s=1}^{\infty}\int_{\mathfrak{r}^1}\frac{\mathfrak{o}_{s1}\phi_s(r)}{\lambda_s}\int_0^{2\pi}\frac{u(re^{i\varphi},re^{-i\varphi})re^{i\varphi}}{re^{i\varphi}-\zeta}d\varphi dr\right]. \tag{4.13}$$

Proof. By (4.3) (with $k=0$) and (4.12), we have

$$A_{n0}=\frac{1}{2\pi R}\sum_{s=1}^{\infty}\int_{\mathfrak{r}^1}\frac{\mathfrak{o}_{s1}\phi_s(r)}{\lambda_s}\int_0^{2\pi}\frac{u(re^{i\varphi},re^{-i\varphi})}{r^ne^{in\varphi}}d\varphi dr. \tag{4.14}$$

Since $Q(\zeta\mid u)=\sum_{n=0}^{\infty}A_{n0}\zeta^n$, the relation (4.13) follows.

THEOREM 4.3. *Under the hypotheses of Theorem* 4.1, *we have*

$$Q(\zeta\mid u)=\frac{1}{2\pi}\sum_{n=0}^{\infty}(n+1/2)P_n(0)\int_{r=-r_1}^{r=r_1}P_n(r/r_1)dr\int_0^{2\pi}\frac{u(re^{i\varphi},re^{-i\varphi})d\varphi}{r_1(1-\zeta/re^{i\varphi})}, \quad |\zeta|<r, \tag{4.15}$$

where the functions P_n are Legendre polynomials.

Proof. Integrating (4.2) multiplied by $e^{-iq\varphi}$ along the circle $|z|=r$, $r\leqq r_1<R$, of the real plane, we obtain

$$\sum_{\nu=0}^{\infty}A_{\nu+q,\nu}r^{2\nu}=\frac{1}{2\pi r^q}\int_0^{2\pi}\frac{u(re^{i\varphi},re^{-i\varphi})}{e^{iq\varphi}}d\varphi. \tag{4.16}$$

Since $\sum_{\nu=0}^{\infty}A_{\nu+q,\nu}r^{2\nu}$ is a function of bounded variation, it can be developed in a uniformly convergent series

$$\begin{aligned}\sum_{\nu=0}^{\infty}A_{\nu+q,\nu}r^{2\nu}&\\ &=\sum_{\nu=0}^{\infty}(\nu+1/2)P_\nu(r/r_1)\int_{r=-r_1}^{r=r_1}\frac{dr}{2\pi r^q}P_\nu(r/r_1)\int_0^{2\pi}\frac{u(re^{i\varphi},re^{-i\varphi})d\varphi}{r_1e^{iq\varphi}}.\end{aligned} \tag{4.17}$$

Substituting $r=0$ we obtain (4.15) from (4.17).

II. Duality between the theory of analytic functions of one complex variable and the theory of functions of $\mathcal{C}(\mathrm{E})$

5. **An integral formula for functions of $\mathcal{C}(\mathrm{E})$.** In this section we shall develop an analogue of the Cauchy integral formula.

THEOREM 5.1. *For every point Z there exists a function $H(z, \bar{z}; Z) \in \mathcal{C}(\mathrm{E})$, regular in the region $\mathrm{E}[|z| < \infty] - \mathrm{E}[z = Zs, 2 \leq s < \infty]$, such that every function $u(z, \bar{z}) \in \mathcal{C}(\mathrm{E})$ can be represented by the line-integral*

$$u(z, \bar{z}) = (1/2\pi i) \int_{\mathfrak{a}^1} R(Z \mid u) H(z, \bar{z}; Z) dZ. \tag{5.1}$$

$\mathfrak{a}^1$ is an arbitrary rectifiable closed curve, lying in the domain of association of u, and such that the origin lies in its interior.

Proof. By (2.1) and the Cauchy integral formula we have

$$\begin{aligned} u(z, \bar{z}) &= \int_{-1}^{1} \mathrm{E}(z, \bar{z}, t) f((z/2)(1 - t^2))(1 - t^2)^{-1/2} dt \\ &= (1/2\pi i) \int_{-1}^{1} \mathrm{E}(z, \bar{z}, t)(1 - t^2)^{-1/2} \left[\int_{\mathfrak{a}^1} f(Z)/(Z - (z/2)(1 - t^2)) dZ \right] dt \\ &= (1/2\pi i) \int_{\mathfrak{a}^1} f(Z) \left[\int_{-1}^{1} [\mathrm{E}(z, \bar{z}, t)/(Z - (z/2)(1 - t^2))](1 - t^2)^{-1/2} dt \right] dZ. \end{aligned} \tag{5.2}$$

The expression in the bracket belongs to $\mathcal{C}(\mathrm{E})$; designating it by $H(z, \bar{z}; Z)$, we obtain (5.1) by (3.1).

6. **Development in series and approximation in $\mathcal{C}(\mathrm{E})$.** If $h_n(z)$ converges to a limit function $h(z)$ for $z \in \mathfrak{F}^2$, $n \to \infty$, then by (2.9),

$$\lim_{n \to \infty} [\mathrm{P}(h_n(z))] = \mathrm{P}(h(z)), \qquad z \in \mathfrak{F}^2. \tag{6.1}$$

This fact enables us to prove, in the theory of functions of $\mathcal{C}(\mathrm{E})$, a large group of theorems dealing with normal families, on development in series, and on approximation.

EXAMPLES. I. Suppose a sequence $u_n(z, \bar{z})$, $n = 1, 2, \cdots$ of functions regular in $\mathfrak{F}^2$ is given, with $u_n \in \mathcal{C}(\mathrm{E}_n)$. Let, furthermore, $\lim_{n \to \infty} \mathrm{E}_n(z, \bar{z}, t) = \mathrm{E}(z, \bar{z}, t)$ for $(z, \bar{z}) \in \mathfrak{F}^2$, $-1 \leq t \leq 1$. Finally, let $Q(z \mid u_n)$ omit (that is, fail to take on) two distinct values. Then $u_n(z, \bar{z})$ form a normal family.

II. As is well known, there exist sets of functions $\{f_\nu(z)\}$ possessing the property that every function f regular in a domain $\mathfrak{F}^2$ can be therein represented in the form $f(z) = \sum a_\nu f_\nu(z)$, where the series converges uniformly in every subdomain $\mathfrak{T}^2$, $\bar{\mathfrak{T}}^2 \subset \mathfrak{F}^2$. To every such theorem corresponds the following analogue: For the domain $\mathfrak{F}^2$ there exists a set of functions $u_n(z, \bar{z})$, $u_n(z, \bar{z}) = \mathrm{P}(f_n(z)) \in \mathcal{C}(\mathrm{E})$, such that every function $u(z, \bar{z}) \in \mathcal{C}(\mathrm{E})$, regular in

$\mathfrak{F}^2$, can be represented in the form $u(z, \bar{z}) = \sum_{\nu=1}^{\infty} a_\nu u_\nu(z, \bar{z})$. This series converges (uniformly) in every $\overline{\mathfrak{T}}^2 \subset \mathfrak{F}^2$. In the same way, every theorem stating that $f(z)$ can be approximated by $\sum_{\nu=1}^{n} a_\nu^{(n)} f_\nu(z)$ in every subdomain $\mathfrak{T}^2$ of $\mathfrak{F}^2$, has an analogue which can be proved in the theory of $\mathcal{C}(\mathrm{E})$. We note that in certain cases, it is possible to approximate $u(z, \bar{z})$ in $\overline{\mathfrak{F}}^2$ (cf. [4]).

The set $\{z^\nu\}$ plays an important role among the sets of functions f, mentioned above. There arises the problem of characterizing the functions $\mathrm{P}(z^{\nu-1})$ independently of their integral representation. This is, in fact, possible if E satisfies a certain differential equation. For then the $\mathrm{P}(z^{\nu-1})$ satisfy an ordinary differential equation. (We shall consider this question in another paper.) In particular, the previous results yield: Every function $u(z, \bar{z}) \in \mathcal{C}(\mathrm{E})$ can be developed in every circle of regularity, $|z| < \rho$, in the form $\sum_{\nu=1}^{\infty} a_\nu \mathrm{P}(z^{\nu-1})$ and it can be approximated by $\sum_{\nu=1}^{n} a_\nu^{(n)} \mathrm{P}(z^{\nu-1})$ in every regularity domain $\mathfrak{F}^2$.

In addition, our method enables us to prove immediately many other theorems concerning the degree of approximation. For instance: Let $w = d(z)$ map conformally the complement of $\mathfrak{F}^2$ into $|w| > 1$, and denote by $\mathfrak{C}_R^1$ the curve $d(z) = R > 1$. If $u(z, \bar{z})$ is analytic in a domain $\mathfrak{F}^2 \supset \mathfrak{C}_R^1$, then there exist expressions $p_n(z, \bar{z}) = \sum_{\nu=1}^{n} a_\nu^{(n)} \mathrm{P}(z^{\nu-1})$ such that

$$\limsup_{n\to\infty} \left[\max \mid u(z, \bar{z}) - p_n(z, \bar{z}) \mid^{1/n}\right] = 1/R.$$

This result is an immediate generalization of the corresponding theorem of Walsh [12].

7. **Coefficient problems.** In an analogous way other results (for instance, those on overconvergence, on existence of boundary values, various gap theorems, and so on) can be proved in the theory of functions of the class $\mathcal{C}(\mathrm{E})$. In §6, we introduced the system $\mathrm{P}(z^{\nu-1})\nu = 1, 2, \cdots$. We indicated that the series

$$\sum_{\nu=1}^{\infty} a_\nu \mathrm{P}(z^{\nu-1}) \tag{7.1}$$

has a behavior analogous to that of a power series in the case of analytic functions of one complex variable. In particular, one can deduce various properties of $u(z, \bar{z})$ from the behavior of the coefficients a_ν of its expansion (7.1). On the other hand, the function $u(z, \bar{z})$ can be represented in the neighborhood of the origin in the form

$$\sum_{\nu=0}^{\infty} \sum_{\mu=1}^{\infty} A_{\nu\mu} z^\nu \bar{z}^\mu, \quad \text{or} \quad \sum_{\nu=0}^{\infty} \sum_{\mu=0}^{\infty} B_{\nu\mu} x^\nu y^\mu, \tag{7.2}$$

the series converging in $\mathrm{E}[|z| < \rho, |\bar{z}| < \rho]$, ρ sufficiently small. The problem now arises of finding properties of u from the behavior of $A_{\nu\mu}$ or $B_{\nu\mu}$.

By (3.4) we have the relation

$$a_m = c_m A_{m0}, \qquad c_m = 2^m \Gamma(m+1)/\pi^{1/2}\Gamma(m+1/2) \tag{7.3}$$

for the coefficient a_m of the function f, which is the associate of u. Thus, if some property of A_{m0} is known, the corresponding property of a_m follows by (7.3). Then, using the theorems of the theory of analytic functions of one complex variable, which deal with the relation between the function and the coefficients of its series development, we may obtain results concerning the relation between the function $u(z, \bar{z})$ and the coefficients A_{m0} of its development (7.2).

EXAMPLES. I. The radius, r, of the largest circle with center at the origin, inside which $u(z, \bar{z}) = \sum A_{mn} z^m \bar{z}^n$ is regular, is given by

$$1/r = \lim_{n\to\infty} (|A_{n0}| c_n)^{1/n}. \tag{7.4}$$

II. Suppose now that $A_{n0} = 0$ in (7.2) for all n, except for $n = \lambda_\nu$, $\nu = 1, 2, \cdots$, where $\lambda_{\nu+1} - \lambda_\nu > \lambda_\nu \theta$, $\theta > 1$. Then $u(z, \bar{z})$ cannot be continued analytically to the outside of the circle whose radius is given by (7.4).

III. A classical result of the theory of entire functions states: Let $f(z) = \sum a_n z^n$ be an entire function. The logarithm of the greatest of the terms $|a_n r^n|$ is asymptotically equal to $\log[\max_{0\le\varphi\le 2\pi} |f(re^{i\varphi})|]$. A similar result is valid in the case of entire functions $u(z, \bar{z}) \in \mathcal{C}(\mathrm{E})$. Namely, we have the inequality

$$|u(re^{i\varphi}, re^{-i\varphi})| \le |A_{n0}R^n|_{\max} \cdot \mathrm{E}_{\max}(r) \cdot (R/R - r), \tag{7.5}$$

where

$$r < R, \ |A_{n0}R^n|_{\max} = \max_{n=0,1,2,\cdots} |A_{n0}R^n|, \ \mathrm{E}_{\max}(r) = \max_{|z|\le r; -1\le t\le 1} |\mathrm{E}(z, \bar{z}, t)|.$$

For we have

$$\begin{aligned} |u(z, \bar{z})| &\le \int_{-1}^{+1} |\mathrm{E}(z, \bar{z}, t)| \cdot \sum_{n=0}^{\infty} |a_n z^n| (1/2)^n (1-t^2)^n (1-t^2)^{-1/2} dt \\ &\le \mathrm{E}_{\max}(r) \cdot \sum_{n=0}^{\infty} |A_{n0} r^n| \le \mathrm{E}_{\max}(r) \cdot |A_{n0}R^n|_{\max} \cdot (R/R - r). \end{aligned}$$

An inequality for A_{n0} in terms of $\max_{0\le\varphi\le 2\pi} |u(re^{i\varphi}, re^{-i\varphi})|$ follows from (4.14).

The relation (7.3) enables us to give interesting formulations of many theorems which have analogues in the theory of functions of $\mathcal{C}(\mathrm{E})$. For instance a generalization of a theorem of Fatou type was given in [5](11). Since the coefficients of the associate function can be expressed by A_{n0} in the form (7.3), it follows from [5]: If $\sum_{m=0}^{\infty} |A_{m0}|^2 < \infty$, then $u = \sum A_{mn} z^m \bar{z}^n = \mathrm{P}(f)$, $u \in \mathcal{C}(\mathrm{E})$, possesses boundary values almost everywhere on the unit

(11) We note that on p. 668, in 1.14 of [5] it is necessary to add to $1 + z\bar{z}t^2\mathrm{E}^*$ the factor $\exp(-\int_0^{\bar{z}} A\, d\bar{z})$.

circle. Further, the set of points in which these boundary values exist, includes the set $\mathfrak{C}$, $\mathfrak{C}$ being the set of points in which $\int_{-1}^{1} f((z/2)(1-t^2))(1-t^2)^{-1/2}dt$ has boundary values.

A generalization of Hadamard's multiplication theorem was given in [2].

For certain applications it is useful to obtain an expression for A_{mn} in terms of A_{n0}. Such a formula follows from the fact that

$$u(z, \bar{z}) = u(z_1 + iz_2, z_1 - iz_2)$$

is an analytic function of two complex variables. Thus, we have

$$\begin{aligned} A_{mn} = & -\frac{1}{4\pi^2 n! m!}\int_0^{2\pi}\int_0^{2\pi}\int_{-1}^{1}\{E[r_1e^{i\varphi_1}+ir_2e^{i\varphi_2},\ r_1e^{i\varphi_1}-r_2e^{i\varphi_2},\ t] \\ & \times \sum_{n=0}^{\infty} A_{n0}[(1/2)(r_1e^{i\varphi_1}+ir_2e^{i\varphi_2})(1-t^2)]^n / r_1^m r_2^n e^{im\varphi_1}e^{in\varphi_2}(1-t^2)^{1/2}\}\, d\varphi_1 d\varphi_2 dt. \end{aligned} \tag{7.6}$$

Using the general integral formula (see [6]), we obtain integral formulas with various ranges of integration.

III. Some problems involving functions of class $\mathcal{C}$(E)

8. **Conjugate functions. Mapping by the functions of $\mathcal{C}$(E).** The functions u considered in this paper are complex. In most applications (theory of linear partial differential equations) we need consider only their real part. We wish to indicate a problem which involves both the real and imaginary part of u.

The equation (1.1) is equivalent to the system of two equations

$$\begin{aligned} &(1/4)\Delta U^{(1)} + (1/2)AU_x^{(1)} + (1/2)BU_y^{(1)} + (1/2)CU_x^{(2)} + (1/2)DU_y^{(2)} \\ &\qquad + c_1U^{(1)} - c_2U^{(2)} = 0, \\ &(1/4)\Delta U^{(2)} - (1/2)CU_x^{(1)} - (1/2)DU_y^{(1)} + (1/2)AU_x^{(2)} + (1/2)BU_y^{(2)} \\ &\qquad + c_2U^{(1)} + c_1U^{(2)} = 0 \end{aligned} \tag{8.1}$$

where

$$U = U^{(1)} + iU^{(2)}, \qquad c = c_1 + ic_2, \qquad A = (1/2)[(a + \bar{a}) + (b + \bar{b})],$$
$$B = (1/2i)[(a - \bar{a}) - (b - \bar{b})], \qquad C = (-1/2i)[(a - \bar{a}) + (b - \bar{b})],$$
$$D = (1/2)[(a + \bar{a}) - (b + \bar{b})].$$

On the other hand, every solution of (1.1) can be written in the form, $\exp(-\int_0^{\bar{z}} a d\bar{z})\cdot u(z, \bar{z}) + \exp(-\int_0^z b dz)\cdot v(z, \bar{z})$, where $v(z, \bar{z})$ belongs to a class of functions, whose associates are anti-analytic functions (that is, analytic functions of $\bar{z}$). Thus the functions of $\mathcal{C}$(E), with an appropriate E, form a subclass of the functions satisfying the system (8.1). However, if $a = \bar{b}$, and c is real, then the equations (8.1) are independent of each other.

Relations existing between the real and imaginary parts of u, in the general case, are given in

THEOREM 8.1. *Let* $u=u^{(1)}+iu^{(2)}\in\mathcal{C}(\mathrm{E})$, $\mathrm{E}=\mathrm{E}^{(1)}+i\mathrm{E}^{(2)}$. *Further, let* $T(z|u^{(1)})+iC(z|u^{(1)})$ *be the associate of* u *(see (4.5)). Then*

$$(8.2)\qquad \begin{aligned} u_x^{(1)} - u_y^{(2)} &= \int_{-1}^{1} (\mathrm{E}_x^{(1)} - \mathrm{E}_y^{(2)})T(z\,|\,u^{(1)})(1-t^2)^{-1/2}dt \\ &\quad - \int_{-1}^{1} (\mathrm{E}_x^{(2)} + \mathrm{E}_y^{(1)})C(z\,|\,u^{(1)})(1-t^2)^{-1/2}dt, \end{aligned}$$

$$(8.3)\qquad \begin{aligned} u_x^{(2)} + u_y^{(1)} &= \int_{-1}^{1} (\mathrm{E}_x^{(2)} + \mathrm{E}_y^{(1)})T(z\,|\,u^{(1)})(1-t^2)^{-1/2}dt \\ &\quad + \int_{-1}^{1} (\mathrm{E}_x^{(1)} - \mathrm{E}_y^{(2)})C(z\,|\,u^{(1)})(1-t^2)^{-1/2}dt. \end{aligned}$$

We obtain (8.2) and (8.3) by differentiating u and using the Cauchy-Riemann differential equations for the associate.

In addition to $u^{(1)}$ and $u^{(2)}$ we may consider the pair of functions $v^{(1)}$, $v^{(2)}$, where $v^{(1)}+iv^{(2)}=\int_{-1}^{1}\mathrm{E}_1(z,\bar{z},t)f[(z/2)(1-t^2)](1-t^2)^{-1/2}dt$ and

$$\mathrm{E}_1=(\mathrm{E}_x^{(1)}-\mathrm{E}_y^{(2)})+i(\mathrm{E}_x^{(2)}+\mathrm{E}_y^{(1)}).$$

The functions $u^{(k)}$ and $v^{(k)}$, $k=1, 2$, are connected by the equations

$$(8.4)\qquad u_x^{(1)} - u_y^{(2)} = v^{(1)}, \qquad u_y^{(1)} + u_x^{(2)} = v^{(2)}.$$

It follows that many relations exist between $u^{(k)}$ and $v^{(k)}$. For instance, if $u^{(1)}$ satisfies a (self-adjoint) partial differential equation of elliptic type and second order, say $L_1(u^{(1)})=\Delta u^{(1)}+4cu^{(1)}=0$, c real, then a generalized Cauchy formula is valid. It determines the values of $u^{(1)}$ inside a domain $\mathfrak{A}^2$ in terms of the values of $u^{(k)}$ and $v^{(k)}$, $k=1, 2$, on the boundary $\mathfrak{a}^1$ of $\mathfrak{A}^2$. Using the formula (9), p. 515 of [11] we obtain

$$\begin{aligned} 2\pi u^{(1)}(x, y) = \int_{\mathfrak{a}^1} [&(u^{(2)}W_\xi - u^{(1)}W_\eta + v^{(2)}W)d\xi \\ &+ (u^{(2)}W_\eta + u^{(1)}W_\xi - v^{(1)}W)d\eta], \qquad (x, y) \in \mathfrak{A}^2. \end{aligned}$$

Here $W(x, y; \xi, \eta)$ is a fundamental solution of L_1.

REMARK. Clearly, if $u^{(k)}$ are connected by the generalized Cauchy-Riemann equations,

$$\sum_{k=1}^{2} [\overset{(s)}{a}_{1k} \overset{(k)}{u}_x + \overset{(s)}{a}_{2k} \overset{(k)}{u}_y + \overset{(s)}{a}_{3k} \overset{(k)}{u} + \overset{(s)}{a}_{4k}] = 0, \qquad s = 1, 2,$$

$$\overset{(1)}{a}_{s1}\, \overset{(2)}{a}_{s2} - \overset{(2)}{a}_{s1}\, \overset{(1)}{a}_{s2} \neq 0,$$

a generalized Cauchy formula can be obtained without introducing $v^{(1)}, v^{(2)}$. If $L_1(u^{(1)})=0$ then

$$2\pi u^{(1)}(x, y) = \int_{\mathfrak{a}^1} \{ [(-(A_2W)_\xi + WC_2 - W_\eta)u^{(1)} + (-(B_2W)_\xi + D_2W)u^{(2)} + E_2W]d\xi - [(-(WA_1)_\eta + WC_1 - W_\xi)u^{(1)} + (-(WB_1)_\eta + WD_1)u^{(2)} + WE_1]d\eta \},$$

$$(x, y) \in \mathfrak{A}^2.$$

$A_k, B_k, \cdots$ are polynomials in $a_{m}^{(k)}$, they are the coefficients of the expressions

$$u_\xi^{(1)} = A_1u_\eta^{(1)} + B_1u_\eta^{(2)} + C_1u^{(1)} + D_1u^{(2)} + E_1,$$
$$u_\eta^{(l)} = A_2u_\xi^{(1)} + B_2u_\xi^{(2)} + \cdots .$$

In analogy with conformal transformations, one may consider the mapping of the (x,y)-plane by the functions $U(z,\bar{z})$ of the class $\exp(-\int_0^{\bar{z}} a d\bar{z})\cdot \mathcal{C}(\mathrm{E})$. If U satisfies (1.1) then this mapping represents a transformation by a pair $(U^{(1)}, U^{(2)})$ of solutions of the system (8.1). The following case is of special interest. Suppose that the boundary $\mathfrak{f}^1$ of $\mathfrak{F}^2$ can be decomposed: $\mathfrak{f}^1=\mathfrak{S}_{\nu=1}^{n}\mathfrak{f}_\nu^1$. Suppose further, that by the transformation $U=U(z, \bar{z})$ every curve-segment $\mathfrak{f}_\nu^1$ is transformed into $\mathfrak{g}_\nu^1=\mathrm{E}[\psi_\nu(U^{(1)}, U^{(2)})=0]$, $\nu=1, 2, \cdots, n$. The pair $(U^{(1)}, U^{(2)})$ then represents a solution of the system (8.2), (8.3), satisfying the boundary condition

$$\psi_\nu(U^{(1)}, U^{(2)})=0 \text{ on } \mathfrak{f}_\nu^1, \qquad \nu=1, 2, \cdots, n.$$

9. **The coincidence of functions of different classes along curves.** In this section we investigate the problem: when can two functions of different classes, or at least their real parts, coincide along a curve. Results in this direction are especially of interest if one of the two classes is the class of analytic functions of one complex variable.

We shall indicate some applications of the results in this direction to the boundary value problem[12], and to the characterization of singularities.

Let $\mathrm{E}(z, \bar{z}, t)$ *possess the property that for* (x, y), $z=x+iy$, $\bar{z}=x-iy$, *belonging to a curve* $\mathfrak{l}^1$ *of the real plane, we have*

$$\mathrm{E}(z, \bar{z}, t) = \mathrm{E}_1(z, t), \qquad (x, y) \in \mathfrak{l}^1, \tag{9.1}$$

---

[12] In analogy with the theory of partial differential equations, we may consider the boundary value problems for the functions of the class $\mathcal{C}(\mathrm{E})$. Since a function which satisfies (1.1) can be represented in the form (2.4), the boundary value problem for $L(U)=0$ can be reduced to that of functions of the class $\mathcal{C}(\mathrm{E})$. We note that if $a=b$ and c is real (see (1.1)), we may write $U(z, \bar{z})=\mathrm{Re}\{\exp(-\int_0^{\bar{z}} a d\bar{z})\int_{-1}^{1}\mathrm{E}(z, \bar{z}, t)f((z/2)(1-t^2))(1-t^2)^{-1/2}dt\}$.

In this case our later results can be directly used in the theory of partial differential equations.

where $E_1(z, t)$ *is an analytic function of one complex variable regular in a* (*sufficiently large*) *domain of the* (*complex*) *z-plane. Then*

$$u(z, \bar{z}) = h(z) \qquad \text{for } (x, y) \in \mathfrak{l}^1 \tag{9.2}$$

where $u(z, \bar{z})$ *is the function of* $\mathcal{C}(E)$ *introduced by* (2.6) *and*

$$h(z) = \int_{-1}^{1} E_1(z, t) f\left(z \frac{1 - t^2}{2}\right) \frac{dt}{(1 - t^2)^{1/2}} \tag{9.3}$$

is an analytic function of one complex variable.

EXAMPLE I. We have

$$E(z, \bar{z}, t) = E(z, 2ic + z, t) \equiv E_1(z, t) \quad \text{for } (x, y) \in \mathfrak{l}_1^1 = E[y = -c]. \tag{9.4}$$

EXAMPLE II. Suppose that

$$E(z, \bar{z}, t) = E_1(r, t) = 1 + t^2 E_1^*(r, t), \qquad r^2 = x^2 + y^2, \tag{9.5}$$

where $E_1(r, t)$ is a function which is independent of φ.

(See also (2.3).) Here r and φ are polar coordinates. Then we have

$$E(z, \bar{z}, t) = E_1(\rho, t) \qquad \text{for } (x, y) \in \mathfrak{l}_2^1 = E[x^2 + y^2 = \rho^2]. \tag{9.6}$$

We now shall discuss the above mentioned applications of the coincidence of the functions u and h on $\mathfrak{l}^1$ (see (9.2)).

1. *Boundary value problem.* We consider at first the case where E is of the form described in Example II. Let $u(z, \bar{z}) \in \mathcal{C}(E)$, where E satisfies (9.6). If for all integers n, $n \geq 0$,

$$J_n(\rho) \equiv \int_{-1}^{1} E_1(\rho, t)(1 - t^2)^{n-1/2} dt \neq 0,$$

then $\tilde{u}(r, \varphi) \equiv u(re^{i\varphi}, re^{-i\varphi})$ can be represented in the domain $\mathfrak{K}^2 = E[|z| < \rho]$ in the form

$$\begin{aligned} \tilde{u}(r, \varphi) &= \frac{1}{2\pi} \int_0^{2\pi} v(\vartheta) \sum_{n=0}^{\infty} \frac{r^n J_n(r) e^{in(\varphi - \vartheta)} d\vartheta}{\rho^n J_n(\rho)} \\ &= \frac{1}{2\pi} \int_0^{2\pi} \int_{-1}^{1} v(\vartheta) E_1(r, \tau) \sum_{n=0}^{\infty} \frac{r^n (1 - \tau^2)^{n-1/2} e^{in(\varphi - \vartheta)} d\vartheta d\tau}{\rho^n J_n(\rho)} \\ &= \frac{1}{(2\pi)^2} \int_0^{2\pi} \int_{-1}^{1} \int_0^{2\pi} \frac{v(\vartheta) E_1(r, \tau) H(be^{-i\alpha})}{1 - rb^{-1}\rho^{-1}(1 - \tau^2) e^{i(\alpha + \varphi - \vartheta)}} \frac{d\tau d\alpha d\vartheta}{(1 - \tau^2)^{1/2}}, \end{aligned} \tag{9.7}$$

where $v(\vartheta) = \tilde{u}(\rho, \vartheta)$ is supposed to be an absolutely integrable function, $r/\rho < b < 1$, and

$$H(be^{-i\alpha}) = \sum_{n=0}^{\infty} \frac{b^n e^{-in\alpha}}{J_n(\rho)}. \tag{9.8}$$

We obtain (9.7) by formal calculation, since it follows from Lemma 9.1 that the first series converges uniformly for $r \leq \rho_0 < \rho$.

LEMMA 9.1. *For every ϵ, $\epsilon > 0$, and every ρ, $\rho \leq \rho_1 < \infty$ there exists an n_0, such that for $n > n_0$ we have*

$$1 - \epsilon \leq \left| \int_{-1}^{1} \mathrm{E}_1(\rho, t)(1 - t^2)^{n-1/2} dt \Big/ \int_{-1}^{1} (1 - t^2)^{n-1/2} dt \right| \leq 1 + \epsilon,$$

$$\int_{-1}^{1} (1 - t^2)^{n-1/2} dt = \frac{\pi^{1/2}\Gamma(n + 1/2)}{\Gamma(n + 1)}. \tag{9.9}$$

Proof. By (9.5) we have

$$\int_{-1}^{1} \mathrm{E}_1(\rho, t)(1 - t^2)^{n-1/2} dt = \int_{-1}^{1} (1 - t^2)^{n-1/2} dt + \int_{-1}^{+1} t^2 \mathrm{E}_1^*(\rho, t)(1 - t^2)^{n-1/2} dt$$

$$= \frac{\Gamma(1/2)\Gamma(n + 1/2)}{\Gamma(n + 1)} + \int_{-1}^{1} t^2 \mathrm{E}_1^*(\rho, t)(1 - t^2)^{n-1/2} dt.$$

(See [10, p. 133 formula (2)].) Since $\mathrm{E}_1^*(\rho, t)$ is supposed to be regular, there exists a constant c, such that $|\mathrm{E}_1^*(\rho, t)| \leq c$, and therefore

$$\left| \int_{-1}^{1} t^2 \mathrm{E}_1^*(\rho, t)(1 - t^2)^{n-1/2} dt \right| \leq c \int_{-1}^{+1} t^2 (1 - t^2)^{n-1/2} dt$$

$$= \frac{c}{2} \frac{\Gamma(1/2)\Gamma(n + 1/2)}{\Gamma(n + 2)}.$$

Hence from $\Gamma(n+2) = (n+1)\Gamma(n+1)$ we have

$$1 - \frac{c}{2(n + 1)} \leq \left| \int_{-1}^{1} \mathrm{E}_1(\rho, t)(1 - t^2)^{n-1/2} dt \Big/ \frac{\Gamma(1/2)\Gamma(n + 1/2)}{\Gamma(n + 1)} \right|$$

$$\leq 1 + \frac{c}{2(n + 1)}$$

which yields (9.9).

Since in (9.7) we supposed $r \leq \rho_0 < \rho$ the absolute and uniform convergence of the series in the first integral of (9.7) follows from (9.9).

An analogous formula can be derived if the derivative $\partial \bar{u}(r, \varphi)/\partial r$ is prescribed along $\mathfrak{k}_2^1$.

REMARK 1. If a (real) function satisfying $L(U) = 0$ and assuming given real values on the boundary has to be determined, we take for $v(\varphi)$ such an analytic function of one complex variable that $\mathrm{Re}[\exp(-\int_0^{\bar{z}} a d\bar{z}) \cdot v(z)]$ assumes the given values of $\mathfrak{k}_2^1$.

REMARK 2. In our above considerations, the existence of functions $u(z, \bar{z}) \in \mathcal{C}(\mathrm{E})$ satisfying the required boundary conditions was presupposed. However, there exist cases where the proof can be given without a preliminary

existence hypothesis([12]). For instance, suppose that a (real) function $H(\varphi)$ is chosen in such a way that $H(\varphi)=\text{Re}[h(\rho e^{i\varphi})]$, where $h(z)$ is an analytic function of one complex variable such that $\int_0^{2\pi}|h(\lambda e^{i\varphi})|^p d\varphi<\infty$, $p>1$.

Let $\sum(a_n\cos n\varphi+b_n\sin n\varphi)$ be the Fourier development of $H(\varphi)$. Then

$$\sum c_n r^n\rho^{-n}e^{in\varphi},\qquad c_n=a_n-ib_n,$$

will be an analytic function, regular in $\mathfrak{K}^2=\text{E}[|z|<\rho]$ the real part of which converges almost everywhere to $H(\varphi)$ when we approach $\mathfrak{k}_2^1$ radially. It follows by Lemma 9.1 that

$$\sum c_n r^n J_n(r)e^{in\varphi}/\rho^n J_n(\rho) \tag{9.10}$$

converges uniformly for every $r\leqq\rho_0<\rho$. Thus (9.10) is a function of $\mathfrak{C}(\text{E})$ which is regular in $\mathfrak{K}^2$, and it suffices to show that it converges to $h(\rho e^{i\varphi})$ as $r\to\rho$. We shall show that

$$\left|\left[\sum_{n=n_0}^{\infty}c_n\frac{r^n}{\rho^n}e^{in\varphi}-\sum_{n=n_0}^{\infty}c_n\frac{r^nJ_n(r)}{\rho^nJ_n(\rho)}e^{in\varphi}\right]\right| \tag{9.11}$$

converges (uniformly in r) to zero, as $n_0\to\infty$. We have

$$\begin{aligned}&\left|\sum_{n=n_0}^{m}c_n\frac{r^n}{\rho^n}e^{in\varphi}-\sum_{n=n_0}^{m}c_n\frac{r^nJ_n(r)}{\rho^nJ_n(\rho)}e^{in\varphi}\right|\\ &\leqq\sum_{n=n_0}^{m}\left|c_n\frac{r^n}{\rho^n}\right|\cdot\left|\frac{J_n(\rho)-J_n(r)}{J_n(\rho)}\right|\\ &\leqq\sum_{n=n_0}^{m}\left|c_n\frac{r^n}{\rho_n}\right|\cdot\left|\int_{-1}^{1}[\text{E}_1(\rho,t)-\text{E}_1(r,t)](1-t^2)^{n-1/2}dt\right|\Big/\int_{-1}^{+1}\text{E}_1(\rho,t)(1-t^2)^{n-1/2}dt.\end{aligned}$$

Since $\text{E}_1(\rho,t)$ is an analytic function of the real variable ρ, it satisfies, for all r, a Lipschitz condition

$$|\text{E}_1(\rho,t)-\text{E}_1(r,t)|\leqq C_1|\rho-r|$$

where C_1 is a fixed constant.

Thus (9.11) is smaller than

$$\sum_{n=n_0}^{m}C_1|\rho-r|\left[\int_{-1}^{1}(1-t^2)^{n-1/2}dt\Big/\int_{-1}^{1}\text{E}_1(\rho,t)(1-t^2)^{n-1/2}dt\right]\cdot|c_n|\frac{r^n}{\rho^n}.$$

([12]) In this case we thus obtain the proof of the existence of a function $u(z,\bar z)\in\mathfrak{C}(\text{E})$ regular in $\text{E}[|z|<\rho]$, the real part of which assumes the prescribed values almost everywhere on the boundary $\text{E}[|z|=\rho]$. (We therefore obtain, in certain instances, an existence proof for solutions of partial differential equations.)

By a known result $\sum_{n=n_0}^{\infty}|c_n|r^n\rho^{-n} \leq \epsilon_1(n_0)/(\rho-r)$ where $\lim_{n\to\infty}\epsilon_1(n)=0$. (See [8, pp. 405–408].) Thus, by (9.9), (9.11) is less than

$$\epsilon_1(n_0)[1+\epsilon(n_0)] = \epsilon_2(n_0), \qquad \lim_{n\to\infty} \epsilon_2(n) = 0,$$

This completes the proof.

In the last paragraph of §8, we considered, a function $u(z, \bar{z}) = u^{(1)} + iu^{(2)}$ which maps a domain $\mathfrak{F}^2 + \mathfrak{f}^1$ into the domain $\mathfrak{G}^2 + \mathfrak{g}^1$, $(\mathfrak{f}^1 = \mathfrak{S}_{\nu=1}^{n}\mathfrak{f}_\nu^1, \mathfrak{g}^1 = \mathfrak{S}_{\nu=1}^{n}\mathfrak{g}_\nu^1)$. The functions $u^{(1)}$ and $u^{(2)}$ are solutions of a system of partial differential equations for which the boundary conditions are: $\psi_\nu[u^{(1)}, u^{(2)}] = 0$ on $\mathfrak{f}_\nu^1$, $\nu = 1, 2, \cdots, n$.

If $\mathfrak{F}^2 = \mathfrak{K}_2^2 = \mathrm{E}[|z| < \rho]$ and if $\psi_\nu(u^{(1)}, u^{(2)})$ are linear functions of $u^{(1)}$ and $u^{(2)}$, the solution of the above boundary value problem for a pair of harmonic functions can be written (in special cases)([14]) in the form $u^{(1)} + iu^{(2)} = c\int_0^z \prod_{\nu=1}^{n}(z-a_\nu)^{\mu_\nu}dz$. This result can be generalized for the functions $u(z, \bar{z}) \in \mathfrak{C}(\mathrm{E})$, E satisfying (9.5). In fact, since by (9.2) $h(z)$ and $u(z, \bar{z})$ coincide on $\mathfrak{k}_2^1$, the determination of $u(z, \bar{z})$ can be reduced to the finding of a function $f(z)$ which satisfies the integral equation

$$(9.12) \qquad \int_{-1}^{1} \mathrm{E}_1(\rho, t) f\left(z\frac{1-t^2}{2}\right)\frac{dt}{(1-t^2)^{1/2}} = h(z), \qquad z \in \mathfrak{K}_2^2,$$

$h(z)$ being the analytic function satisfying the prescribed boundary conditions. We can develop $f(z)$ and $h(z)$ in power series in the domain $\mathfrak{K}_2^2$. Writing $h(z) = \sum_{n=0}^{\infty} a_n z^n$ and comparing coefficients, we get

$$(9.13) \qquad f(\zeta) = \sum_{n=0}^{\infty} \frac{2^n a_n \zeta^n}{J_n(\rho)} = \frac{1}{2\pi}\int_0^{2\pi} h\left(\frac{2\zeta}{b}e^{i\alpha}\right)\mathrm{H}(be^{-i\alpha})d\alpha,$$

where H is the function introduced in (9.8), and $|\zeta| < b < 1$. Since in the case considered we have $h(z) = \int_0^z \prod_{\nu=1}^{n}(z-a_\nu)^{\mu_\nu}dz$, we have

$$(9.14) \qquad \begin{aligned} u(z, \bar{z}) = \frac{1}{2\pi}\int_{-1}^{1}\int_0^{2\pi}\int_0^{s} \mathrm{E}(z, \bar{z}, t)\mathrm{H}(be^{-i\alpha}) \\ \cdot \left[\prod_{\nu=1}^{n}(\eta - a_\nu)^{\mu_\nu}\right](1-t^2)^{-1/2}d\eta d\alpha dt, \; s = z(1-t^2)e^{i\alpha}/b. \end{aligned}$$

([14]) We note that in the case considered $u^{(1)}$ and $u^{(2)}$ satisfy the Cauchy-Riemann equations in addition to the potential equation.

Since $u^{(1)} + iu^{(2)} = \exp(\int_0^z a d\bar{z})U$ (see (2.5)) the relations $A_\nu u^{(1)} + B_\nu u^{(2)} + C_\nu = 0$, $\nu = 1, 2, \cdots, n$; A_ν, B_ν, C_ν being constants, become $(A_\nu + B_\nu)p_1 U^{(1)} + (A_\nu - B_\nu)p_2 U^{(2)} + C_\nu = 0$ where p_1 and p_2 are the real and imaginary parts, respectively, of $\exp(\int_0^z a d\bar{z})$.

We remark that when dealing with differential equations, especially in connection with the coincidence problem, it is often useful to consider classes $\mathfrak{C}(\mathrm{E})$ with a generating function E which does not fulfill the hypothesis A (see p. 133).

In the case I (see p. 146) we can proceed similarly. However, the determination of f from (9.3) is slightly more complicated. Let $h(z)$ be the analytic function, regular in $\mathfrak{R}_1^2=\mathrm{E}[y>-c]$, $c>0$, which assumes the given values on the curve $\mathfrak{l}_1^1=\mathrm{E}[y=-c]$. Since it is a convex domain containing the origin, there exists, by Corollary 3.1, an analytic function $f(z)=R(z|u)$ (see (3.1)) such that

$$\int_{-1}^{1}\mathrm{E}(z, z+2ic, t)f\left(z\frac{1-t^2}{2}\right)\frac{dt}{(1-t^2)^{1/2}} = h(z). \tag{9.15}$$

Let $\sum\alpha_n z^n$ and $\sum A_n z^n$ be the function elements of f and h, respectively, at the origin and suppose $\mathrm{E}(z, z+2ic, t)=\sum \mathrm{P}_n(t)z^n$. Then we have

$$\alpha_n = \left|\begin{matrix} E_{0,0} & 0 & 0 & \cdots A_0 \\ E_{1,0} & E_{0,1} & 0 & \cdots A_1 \\ \cdots & \cdots & \cdots & \cdots \\ E_{n,0} & E_{n-1,1} & E_{n-2,2} & \cdots A_n \end{matrix}\right| \Bigg/ \prod_{\nu=0}^{n} E_{0,\nu}, \tag{9.16}$$

$$E_{n,k} = \int_{-1}^{1}\mathrm{H}_n(t)(1-t^2)^{k-1/2}dt/2^k.$$

Proof. A formal calculation yields,

$$\int_{-1}^{1}\left\{\sum_{n=0}^{\infty}\mathrm{H}_n(t)z^n\sum_{k=0}^{\infty}\alpha_k z^k 2^{-k}(1-t^2)^{k-1/2}\right\}dt$$

$$= \sum_{n=0}^{\infty}z^n\left[\sum_{\nu=0}^{n}\alpha_\nu\int_{-1}^{1}\mathrm{H}_{n-\nu}(t)2^{-\nu}(1-t^2)^{\nu-1/2}dt\right].$$

By a comparison of coefficients, we have $\sum_{\nu=0}^{n}\alpha_\nu E_{n-\nu,\nu}=A_n$, which yields (9.16).

2. *Residue theorems.* There exists a simple method for the construction of functions of the class $\mathfrak{C}(\mathrm{E})$ with certain singularities In fact suppose that $f(z)$ is a function, which is regular at the origin and possesses a singularity at the point α. For example, take $f(z)=(z-\alpha)^{-1}$. The function[15] $u(z, \bar{z})=\mathrm{P}[(z-\alpha)^{-1}]$ will belong by definition (see §2) to the class $\mathfrak{C}(\mathrm{E})$, and will be defined by the integral representation (2.6) over the domain $\mathfrak{B}^2=\mathrm{E}[|z|<\infty]-\mathrm{E}[z=2\alpha S,\ 1\leqq S<\infty]$. As we shall show immediately, $u(z, \bar{z})$ is also

[15] Sometimes it is useful, for the construction of functions with singularities to use operators slightly different from (2.6), for example, operators of the form $\mathrm{P}_1(f)=\mathrm{E}_1(z, \bar{z})f(z)+\int\mathrm{E}_2(z, \bar{z}, t)f[p(z, t)]dt$, or involving double integrals. If $f(z)$ becomes infinite in such a way that the above integral is a regular function of $z, \bar{z}$, then $\mathrm{P}_1(f)$ possesses a singularity of the same character as f. (See also §1.) In particular, this method yields fundamental solutions.

Note that by the method indicated in [3, especially, p. 1173] various representations of this form can be easily obtained, E_1 and E_2 being solutions of certain integral or differential equations.

regular on $E[z=2\alpha S,\ 1<S<\infty]$. Let $z^0=Re^{i\varphi}$, $R=2|\alpha|S$, $S>1$, and $\varphi=\arg\alpha$. The function $[z^0((1-t^2)/2)-\alpha]^{-1}(1-t^2)^{-1/2}$ considered as a function of the complex variable $t=t_1+it_2$ possesses two poles, namely at the points $t=t^{(1)}\equiv+(1-S^{-1})^{1/2}$ and $t=t^{(2)}\equiv-(1-S^{-1})^{1/2}$. We write

$$\mathfrak{S}_1^1 = E[-1<t_1<1,\ t_2=0],$$

$$\mathfrak{S}_2^1 = E[-1<t_1<1,\ t_2=0] - \mathop{\mathrm{S}}_{k=1}^{2} E[-\epsilon+t^{(k)}<t_1<t^{(k)}+\epsilon,\ t_2=0] + \mathop{\mathrm{S}}_{k=1}^{2} E[t=t^{(k)}+\epsilon e^{i\varphi},\ 180^\circ \leqq \varphi \leqq 360^\circ],$$

$$\mathfrak{S}_3^1 = E[-1<t_1<1,\ t_2=0] - \mathop{\mathrm{S}}_{k=1}^{2} E[-\epsilon+t^{(k)}<t_1<t^{(k)}+\epsilon,\ t_2=0] + \mathop{\mathrm{S}}_{k=1}^{2} E[t=t^{(k)}+\epsilon e^{i\varphi},\ 0 \leqq \varphi \leqq 180^\circ],$$

ϵ being sufficiently small.

For every $z^{(1)}=Re^{i\psi}$, $\psi<\arg\alpha$, we have

$$\int_{\mathfrak{S}_1^1} [(z/2)(1-t^2)-\alpha]^{-1}(1-t^2)^{-1/2}dt = \int_{\mathfrak{S}_2^1} [(z/2)(1-t^2)-\alpha]^{-1}(1-t^2)^{-1/2}dt$$

since $\mathfrak{S}_1^1$ can be reduced to $\mathfrak{S}_2^1$ without cutting across singularities. The second integral exists, and represents an analytic function even for $z^{(1)}=z^{(0)}$, since the integrand is regular on $\mathfrak{S}_2^1$. It follows that $\int_{-1}^{1}[(z/2)(1-t^2)-\alpha]^{-1}(1-t^2)^{-1/2}dt$ is a regular function of z in the point $z^{(0)}$. However, in general we shall find different values for this function if we approach first from the left and then from the right, since $\mathfrak{S}_3^1$ cannot be reduced to $\mathfrak{S}_2^1$ without cutting the poles at $t^{(1)}$, $t^{(2)}$. Hence the point 2α will in general be a branch point of $P[(z-\alpha)^{-1}]$.

Suppose that $f(z)$ possesses a denumerable number of singularities which have no accumulation point within a finite distance of the origin. Then an analogous consideration shows that $P(f)$ will possess, at corresponding points, singularities which are, in general, branch points of $P(f)$. In that case, the integral formula (2.6) represents one branch of $P(f)$.

Now the problem of characterizing these singularities arises[16]. If the functions $u(z,\bar{z})$ belonging to a class $\mathfrak{C}(E)$ coincide with analytic functions of one complex variable, along certain curves, we may use this fact for one kind of characterization of singularities.

The procedure which can be applied may be demonstrated in the case where E satisfies the relation (9.5).

[16] If $E(z,\bar{z},t)$ satisfies certain differential equations, the function $P[(z-\alpha)^{-1}]$ satisfies *ordinary* differential equations, with coefficients which are connected with $E(z,\bar{z},t)$ in a simple way. For details see [2].

Suppose at first that $u(z, \bar{z})$ is regular in the circle $\mathfrak{K}^2 = \mathrm{E}[|z| \leq \rho]$. Then by Corollary 3.1, $f(z)$, (and hence also, by (9.3), $h(z)$) is regular in $\mathfrak{K}^2$, and we have

$$\int_{\mathfrak{k}^1} u(z, \bar{z})dz = \int_{\mathfrak{k}^1} h(z)dz = 0, \qquad \mathfrak{k}^1 = \mathrm{E}[|z| = \rho]. \tag{9.17}$$

Thus under the conditions indicated above, the line integral (9.17) taken along a *circle* vanishes if $\mathfrak{k}^1$ lies in regularity domain of $u(z, \bar{z})$.

Suppose now that $f(z)$ has a pole, that is, say we have $f(z) = (z-\alpha)^{-1}$ and $\rho > 2\alpha$. Since the point 2α is a branch point, $\mathfrak{k}^1 = \mathrm{E}[|z| = \rho]$ will be now an *open* curve on the Riemann surface of $u(z, \bar{z})$. (It lies in the sheet in which the representation (2.5) is valid.) Both end points of $\mathfrak{k}^1$ lie on the slit

$$\mathrm{E}[z = 2\alpha S,\ 1 \leq S < \infty]$$

(but of course in different sheets of the Riemann surface).

We have, then,

$$\begin{gathered}\int_{\mathfrak{k}^1} \mathrm{P}[(z - \alpha)^{-1}]dz = 4\pi i \int_{t=t^{(2)}}^{t=t^{(1)}} \mathrm{E}_1(\rho, t)(1 - t^2)^{-3/2}dt, \\ t^{(n)} = -(-1)^n(1 - 2|\alpha|\rho^{-1})^{1/2}, \qquad n = 1, 2.\end{gathered} \tag{9.18}$$

Proof. The left-hand member of (9.18) can be written in the form

$$\int_{\mathfrak{k}^1} \int_{-1}^{1} \mathrm{E}_1(\rho, t)[(z/2)(1 - t^2) - \alpha]^{-1}(1 - t^2)^{-1/2}dtdz. \tag{9.19}$$

The integrand of (9.19) is an absolutely integrable function. We have therefore

$$\int_{\mathfrak{k}^1} \int_{t=-1}^{t=1} \cdots = \int_{\mathfrak{k}^1} \int_{t>t^{(2)}}^{t<t^{(1)}} \cdots + \int_{\mathfrak{k}^1} \int_{t>-1}^{t<t^{(2)}} \cdots + \int_{\mathfrak{k}^1} \int_{t>t^{(1)}}^{t<1} \cdots .$$

Changing the order of integration, the residue theorem then gives

$$\int_{\mathfrak{k}^1} \frac{\mathrm{E}_1(\rho, t)dz}{[(z/2)(1 - t^2) - \alpha](1 - t^2)^{1/2}} = \begin{cases} 4\pi i \dfrac{\mathrm{E}_1(\rho, t)}{(1 - t^2)^{3/2}} & \text{for} \quad t^{(2)} < t < t^{(1)}, \\ 0 & \text{for} \quad -1 < t < t^{(2)}, \\ 0 & \text{for} \quad t^{(1)} < t < 1, \end{cases}$$

which yields (9.18).

The analogous formula for $\mathrm{P}[(z-\alpha)^{-n}]$, $n > 1$, an integer, can be obtained in the following way: The integral, in a certain neighborhood of every point $\alpha = \alpha^0$ for which $|\alpha^0| \neq \rho$ is a regular function of α.

Differentiating (9.18) n times with respect to α we obtain

$$(9.20) \qquad (-1)^n n! \int_{\mathfrak{l}^2} \mathrm{P}[(z-\alpha)^{-(n+1)}]dz = 4\pi i \frac{d^n}{d\alpha^n}\left\{\int_{t=t^{(2)}}^{t=t^{(1)}} \frac{\mathrm{E}_1(\rho, t)dt}{(1-t^2)^{1/2}}\right\}.$$

(Note that α appears only in $t^{(1)}$ and $t^{(2)}$.)

10. A connection with a class of difference equations. There exists an important connection between differential and difference equations. In particular, some of our previous results can be used for the theory of difference equations of the type[17]

$$(10.1) \qquad \begin{aligned} &(M+1)(N+1)\psi(M+1, N+1) \\ &\quad + \sum_{S=0}^{S_1}\sum_{K=0}^{K_1} \alpha_{SK}(M-S+1)\psi(M-S+1, N-K) \\ &\quad + \sum_{S=0}^{S_2}\sum_{K=0}^{K_2} \beta_{SK}(N-K+1)\psi(M-S, N-K+1) \\ &\quad + \sum_{S=0}^{S_3}\sum_{K=0}^{K_3} \gamma_{SK}\psi(M-S, N-K) = 0. \end{aligned}$$

Here α_{SK}, β_{SK}, and γ_{SK} are constants.

THEOREM 10.1. *Let*

$$(10.2) \qquad U(z, \bar{z}) = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} \psi(m, n) z^m \bar{z}^n$$

be a solution of (1.1), *where*

$$(10.3) \qquad a = \sum_{s=0}^{S_1}\sum_{k=0}^{K_1} \alpha_{sk} z^s \bar{z}^k, \qquad b = \sum_{s=0}^{S_2}\sum_{k=0}^{K_2} \beta_{sk} z^s \bar{z}^k, \quad \textit{and} \quad c = \sum_{s=0}^{S_3}\sum_{k=0}^{K_3} \gamma_{sk} z^s \bar{z}^k.$$

Then ψ (M, N) *is a solution of the difference equation* (10.1).

Proof. We have

$$U_{z\bar{z}} = \sum_M \sum_N (M+1)(N+1)\psi(M+1, N+1) z^M \bar{z}^N,$$

$$U_z \sum_{S=0}^{S_1}\sum_{K=0}^{K_1} \alpha_{SK} z^S \bar{z}^K = \sum_M \sum_N \sum_{S=0}^{S_1}\sum_{K=0}^{K_1} \alpha_{SK}(M-S+1)\psi(M-S+1, N-K) z^M \bar{z}^N,$$

$$U_{\bar{z}} \sum_{S=0}^{S_2}\sum_{K=0}^{K_2} \beta_{SK} z^S \bar{z}^K = \sum_M \sum_N \sum_{S=0}^{S_2}\sum_{K=0}^{K_2} \beta_{SK}(N-K+1)\psi(M-S, N-K+1) z^M \bar{z}^N,$$

$$U \sum_{S=0}^{S_3}\sum_{K=0}^{K_3} \gamma_{SK} z^S \bar{z}^K = \sum_M \sum_N \sum_{S=0}^{S_3}\sum_{K=0}^{K_3} \gamma_{SK}\psi(M-S, N-K) z^M \bar{z}^N.$$

[17] We note that the difference equations in two variables have been treated very little by analytical methods. As far as I know the only results in this direction were obtained by C. R. Adams, *On the existence of solutions of a linear partial pure difference equation*, Bull. Amer. Math. Soc. vol. 32 (1926) p. 197.

Since U is supposed to satisfy equation (1.1), the function $\psi(M, N)$ must satisfy (10.1).

REMARK. In speaking of a solution, $\psi(M, N)$ of (10.1), we shall in the future always suppose that the $\psi(M, N)$ have the following property: There exists a number $\rho > 0$ such that $\sum_{M=0}^{\infty}\sum_{N=0}^{\infty}|\psi(M, N)|\rho^{M+N} < \infty$. Under this hypothesis, it follows that conversely, to every solution $\psi(M, N)$ of (10.1) there corresponds a function $U(z, \bar{z})$ given by (10.2) and satisfying $L(U)=0$.

The connection indicated in Theorem (10.1) enables one to reduce many problems of the theory of difference equations of type (10.1) to that of functions satisfying $L(U)=0$. Then the application of the methods of the theory of partial differential equations may yield the desired result.

As an example of such a procedure, the following problem may be considered:

To give a representation of the solution $\psi(M, N)$ of (10.1) in terms of $\psi(N, 0)$ and $\psi(0, N)$, $n=0, 1, 2, \cdots$.

By Theorem (10.1) (cf. also the remark following it) this problem is equivalent to finding the coefficients of U satisfying (1.1), where the functions a, b and c are given by (10.3). On the other hand, by (7.6) we have for the coefficients A_{mn} the relation

$$\psi(M, N) =$$

$$(10.4)\quad -\frac{1}{4\pi^2 M!N!}\int_0^{2\pi}\int_0^{2\pi}\int_{-1}^{1}\sum_{k=1}^{2}\mathrm{H}_k\mathrm{E}_k f_k d\phi_1 d\phi_2 dt \Big/ r_1^M r_2^N e^{i(M\phi_1+N\phi_2)}(1-t^2)^{1/2},$$

Here

$$H_1 = \exp\left[-\int_0^{(r_1e^{i\phi_1}-ir_2e^{i\phi_2})} a(r_1e^{i\phi_1}+ir_2e^{i\phi_2}, \bar{z})d\bar{z}\right],$$

$$H_2 = \exp\left[-\int_0^{(r_1e^{i\phi_1}+ir_2e^{i\phi_2})} b(z, r_1e^{i\phi_1}-ir_2e^{i\phi_2})dz\right],$$

$$f_1 = \sum_{M=0}^{\infty}\frac{\psi(M, 0)}{M!}\left[\frac{(r_1e^{i\phi_1}+ir_2e^{i\phi_2})(1-t^2)}{2}\right]^M,$$

$$f_2 = \sum_{N=0}^{\infty}\frac{\psi(0, N)}{N!}\left[\frac{(r_1e^{i\phi_1}-ir_2e^{i\phi_2})(1-t^2)}{2}\right]^N,$$

and the functions $\mathrm{E}_k = \mathrm{E}_k(r_1e^{i\phi_1}+ir_2e^{i\phi_2},\ r_1e^{i\phi_1}-ir_2e^{i\phi_2},\ t)$, $k=1, 2$, are generating functions of the totality of functions satisfying (1.1) when a, b, c are given by (10.3).

The representation (10.4) enables us to draw various conclusions concerning $\psi(M, N)$. For instance, the growth properties of $\psi(M, N)$ (considered

as a function of M and N), in terms of the growth properties of $\psi(N, 0)$ and $\psi(0, N)$, can be obtained from this relation.

Bibliography

1. Stefan Bergman, *Anwendung des Schwarzschen Lemmas auf eine Klasse von Funktionen von zwei komplexen Veränderlichen*, Jber. Deutschen Math. Verein. vol. 39 (1930) pp. 163–168.

2. ———, *Sur un lien entre la théorie des équations aux derivées partielles elliptiques et celle des fonctions d'une variable complexe*, C. R. Acad. Sci. Paris vol. 205 (1937) pp. 1198–1200 and 1360–1362.

3. ———, *Zur Theorie der Funktionen, die eine lineare partielle Differentialgleichung befriedigen* Rec. Math. (Mat. Sbornik) N.S. vol. 2 (1937) pp. 1169–1198.

4. ———, *The approximation of functions satisfying a linear partial differential equation*, Duke Math. J. vol. 6 (1940) pp. 537–561.

5. ———, *Boundary values of functions satisfying a linear partial differential equation of elliptic type*, Proc. Nat. Acad. Sci. U.S.A. vol. 26 (1940) pp. 668–671.

6. ———, *Ueber eine Integraldarstellung von Funktionen zweier komplexer Veränderlichen*, Rec. Math. (Mat. Sbornik) N.S. vol. 1 (1936) pp. 851–862.

7. ———, *Sur les fonctions orthogonales de plusieurs variables complexes*, Interscience Publishers, New York, 1941.

8. G. H. Hardy and J. E. Littlewood, *Some properties of fractional integrals*. II, Math. Zeit. vol. 34 (1932) pp. 403–439.

9. D. Hilbert, *Grundzüge einer allgemeiner Theorie der linearer Integralgleichungen*, Leipzig and Berlin, 1912.

10. N. Nielsen, *Handbuch der Theorie der Gammafunktionen*, Leipzig, 1906.

11. A. Sommerfeld *Randwertaufgaben in der Theorie der partiellen Differentialgleichungen*. Enzyklopädie der mathematischen Wissenschaften, vol. II, A, 7c pp. 504–570.

12. J. L. Walsh, *Interpolation and approximation by rational functions in the complex domain*, Amer. Math. Soc. Colloquium Publications vol. 20, 1935.

Brown University,
Providence, R.I.

as a function of M and N, in terms of the growth [illegible] can be obtained from this relation.

[illegible]

[illegible]

SOME EINSTEIN SPACES WITH CONFORMALLY SEPARABLE FUNDAMENTAL TENSORS

BY
YUNG-CHOW WONG(1)

1. Introduction. When the fundamental tensor(2) of a Riemannian m-space(3) V_m is of the form (4)

$$(1.1)\qquad {}^*g_{\alpha\beta} = \begin{bmatrix} \rho^{-2}g_{ij} & 0 \\ 0 & \sigma^{-2}g_{pq} \end{bmatrix}, \qquad \begin{aligned} &\alpha, \beta, \gamma, \delta, \epsilon = 1, \cdots, m, \\ &i, j, k, l = 1, \cdots, n, \\ &p, q, r, s = n+1, \cdots, m, \end{aligned}$$

where

$$(1.2)\qquad \begin{aligned} \rho &= \rho(x^\alpha), & \sigma &= \sigma(x^\alpha), \\ g_{ij} &= g_{ij}(x^k), & g_{pq} &= g_{pq}(x^r), \end{aligned}$$

it is said to be *conformally separable* of the type $(n, m-n)$; the tensors ${}^*g_{ij}=\rho^{-2}g_{ij}$ and ${}^*g_{pq}=\sigma^{-2}g_{pq}$, with x^r and x^k, respectively, as parameters, are called its *component tensors*. We shall say that the tensor (1.1) is ***properly*** or ***improperly*** conformally separable according as $\partial_p\rho\neq 0$, $\partial_i\sigma\neq 0$(5) are satisfied or not satisfied.

The tensor (1.1) as a generalization of the ordinarily separable tensor [1, p. 124](6) was recently introduced by Yano [14], where he proved that *in a V_m with fundamental tensor*

$$ {}^*g_{\alpha\beta} = \begin{bmatrix} {}^*g_{ij} & 0 \\ 0 & {}^*g_{pq} \end{bmatrix}, $$

Presented to the Society, October 31, 1942; received by the editors November 24, 1941.

(1) Most of the results in this paper were obtained while I was a Chinese Ying-Keng Funds Student visiting Princeton University, for the courtesy of whose authorities, especially Professor L. P. Eisenhart, I wish to express my most sincere thanks. I wish also to thank Professor D. J. Struik of Massachusetts Institute of Technology for the conversations we had from time to time during the preparation of the manuscript.

(2) Fundamental tensors are always supposed to be nonsingular, though not necessarily definite. All functions appearing in this paper are assumed to have differentiability properties adequate to the part they play in the discussion.

(3) We denote by V, S, E a Riemannian space, a space of constant curvature, and an Einstein space, respectively. The dimensionality is denoted, if necessary, by an index at the lower right-hand corner.

(4) An index has the same range throughout this paper. An index which appears twice in an expression is to be summed over the appropriate range. A free index of a tensor equation assumes each value of its range. A numerical index at the upper right-hand corner of a letter means an exponential, except in the case of the coordinates x^α, x^k or x^r.

(5) We use the notation $\partial_\alpha\equiv\partial/\partial x^\alpha$.

(6) Such a reference is made to the literatures listed at the end of this paper.

$n>1$, *the subspaces* $x^r = const.$ *are totally umbilical, if and only if* $^*g_{ij}$ *is of the form* $[\rho(x^a)]^{-2} g_{ij}(x^k)$. He also proved that *if a conformally separable tensor represents*(7) *an* S_m (that is, an m-space of constant curvature), *then each of its component tensors, if it is of dimension greater than* 2, *represents* S*'s.*

By definition, an Einstein space E is a V whose Ricci and fundamental tensors differ only by a scalar factor(8). The result mentioned at the end of the last paragraph no longer holds if S is replaced by E, although an S is a special E. In this paper we present a complete study of the conformally separable tensor which represents an E_m and each of whose component tensors either is of dimension less than 3 or represents E's. It is found that the construction of such a conformally separable tensor is invariably reduced to that of the fundamental tensor g_{ij} of an E_n or a V_2 for which the following equation admits a solution(9) for y:

$$y_{,ij} = -Ig_{ij}, \tag{1.3}$$

where the comma denotes covariant differentiation with respect to g_{ij}, and I is an unspecified scalar. We shall be content with this result, because the latter problem has already been considered in detail by Brinkmann [2, 3] in his study of E's which are conformal to each other.

In §2, some results concerning the differential equation (1.3) are given. In §3, we find the expressions for the Riemann and Ricci tensors of the tensor (1.1) in terms of those for the same-named tensors of its component tensors. Concerning a properly conformally separable tensor of the type ($n>1$, $m-n=1$), which we consider in §4, we prove (1) that *if an* E_m *admits a one-parameter family of totally umbilical hypersurfaces, then they are conformal to one another and each of them has constant scalar curvature* (Theorem 4.1); and (2) that *a one-parameter family of conformal* E_n*'s with fundamental tensors* $[\rho(x^k, x^m)]^{-2}g_{ij}(x^k)$ *can in general be imbedded isometrically in an* E_{n+1} *as totally umbilical hypersurfaces* (Theorem 4.2). §§5 and 6 are devoted to the study of a properly conformally separable tensor $^*g_{\alpha\beta}$ of the type ($n>1$, $m-n>1$) which represents an E_m and each of whose component tensors is either of dimension 2 or represents E's. By means of Theorem 5.1 on a certain system of differential equations, we show that $^*g_{\alpha\beta}$ *is conformal to an ordinarily separable tensor of the type* (n, $m-n$) (Theorem 5.2). This result enables us to prove that *the component tensors of* $^*g_{\alpha\beta}$ *have the property that, either each of them represents* E*'s or* S_2*'s, or* $n=m-n=2$ *and neither of them represents* S_2*'s* (Theorem 6.2). Characteristic properties of $^*g_{\alpha\beta}$ are then derived (Theorems 6.3 and 6.4), showing how the construction of $^*g_{\alpha\beta}$ depends on that of the

(7) We sometimes find it convenient to express the fact that g_{ij} is the fundamental tensor of an S (or E) by saying that g_{ij} represents an S (or E).

(8) A V_2 is always an E_2, and an E_3 is identical with an S_3 [13]. For convenience, we agree that whenever we speak of an E, it is understood that E is of dimension greater than 2.

(9) By "solution" we always mean non-constant solution.

fundamental tensor of an E_n or a V_2 which admits a solution of (1.3).

The discussion of improperly conformally separable tensors is much easier and is carried out in §§7 and 8. In §9, the theorem of Yano concerning an S_m with conformally separable fundamental tensor is extended, and the paper ends at §10 with some canonical forms for the conformally separable tensors of the type (2, 2) which represents E_4's.

We conclude this introduction with the following remarks. Since the component tensor $^*g_{ij}$ of the tensor (1.1) can be written as

$$^*g_{ij} = \left[\frac{\rho(x^k, x^r)}{\rho(x^k, x_0^r)}\right]^{-2} [\rho(\overset{k}{x}, \overset{r}{x_0})]^{-2} g_{ij},$$

where x_0^r are certain fixed values of x^r, there is no loss of generality in assuming that the function ρ is such that

$$\rho(\overset{k}{x}, \overset{r}{x_0}) = 1; \tag{1.4}$$

in particular, if $\partial_p \rho = 0$, we may assume that $\rho = 1$. This assumption will be made whenever it is desirable. A similar remark holds for the function σ. Finally, the fundamental tensor of every V_2 referred to orthogonal coordinate curves is conformally separable, and for this reason we shall always suppose that $m > 2$.

I. Preliminaries

2. **The equation** $y_{,ij} = -Ig_{ij}$. In what follows we have frequent occasions to meet the following differential equation in the unknown scalar y:

$$y_{,ij} = -Ig_{ij}, \qquad f, h, i, j, k, l = 1, \cdots, n\ (> 1), \tag{2.1}$$

where I is an unspecified scalar and the comma denotes covariant differentiation with respect to the fundamental tensor g_{ij}. This equation has been considered by several authors for different purposes (Brinkmann [3, pp. 121–124]; Fialkow [7, pp. 426–427; 8, pp. 471–473]; Yano [16]; Delgleize [4]). Here we confine ourselves to the case when the Ricci tensor R_{ij} of the V_n with fundamental tensor g_{ij} satisfies

$$R_{ij} = -(n-1)ag_{ij}, \qquad a = \text{a scalar}, \tag{2.2}$$

that is, when V_n is a $V_2(a)$ or an $E_n(a)$[10]. In the latter case, (2.2) implies that a = const. [5, p. 93, Exercise 5].

We first find a geometric meaning of (2.1). The Ricci tensors R_{ij} and $\bar{R}_{ij}$ of g_{ij} and $\bar{g}_{ij} = y^{-2}g_{ij}$ are connected by [5, p. 90, (28.6)]

$$\bar{R}_{ij} = R_{ij} - (n-2)\frac{y_{,ij}}{y} + g_{ij}\left[-\frac{g^{hk}y_{,hk}}{y} + (n-1)\frac{g^{hk}y_{,h}y_{,k}}{y^2}\right]. \tag{2.3}$$

(10) We denote an E or S of scalar curvature a by $E(a)$ or $S(a)$, respectively.

From this it follows at once that

THEOREM 2.1. *Given an E_n $(n>2)$ with fundamental tensor g_{ij}, the V_n with fundamental tensor $y^{-2}g_{ij}$ is also an E_n, if and only if y satisfies* (2.1).

By definition, a V_n is an S_n if the Riemann tensor of V_n is of the form

$$R^{l}_{ijk} = a(\delta^{l}_{j}g_{ik} - \delta^{l}_{k}g_{ij}), \qquad a = \text{const.} \tag{2.4}$$

A consequence of this is that an S_n is necessarily an E_n. The Riemann tensors R^l_{ijk} and $\overline{R}^l_{ijk}$ of g_{ij} and $\bar{g}_{ij}=y^{-2}g_{ij}$ are connected by [5, p. 90, (28.10)]

$$\begin{aligned}\overline{R}^{l}_{ijk} = R^{l}_{ijk} &+ (\delta^{l}_{j}y_{,ik} - \delta^{l}_{k}y_{,ij})y^{-1} + g^{hl}(y_{,hj}g_{ik} - y_{,hk}g_{ij})y^{-1} \\ &- g^{jh}y_{,j}y_{,h}(\delta^{l}_{j}g_{ik} - \delta^{l}_{k}g_{ij})y^{-2}.\end{aligned} \tag{2.5}$$

From (2.4), (2.5) and Theorem 2.1 it can easily be proved that

THEOREM 2.2. *Theorem* 2.1 *remains true when E_n is replaced by S_n.*

We now suppose that (2.1) has a solution y. The integrability condition of (2.1) is

$$R^{l}_{ijk}y_{,l} = y_{,ijk} - y_{,ikj} = - g_{ij}I_{,k} + g_{ik}I_{,j}. \tag{2.6}$$

Transvecting this by the contravariant components g^{ij} of g_{ij}, we have

$$R^{l}_{k}y_{,l} = -(n-1)I_{,k}, \tag{2.7}$$

where $R^l_k = g^{hl}R_{hk}$. When (2.2) is satisfied, (2.7) becomes

$$ay_{,k} = I_{,k}, \tag{2.8}$$

from which it follows that

$$I = I(y), \qquad a = a(y) = dI/dy. \tag{2.9}$$

By differentiating $g^{ij}y_{,i}y_{,j}$ and then making use of (2.1) and (2.9), we obtain

$$(g^{ij}y_{,i}y_{,j})_{,k} = -2I(y)y_{,k},$$

so that

$$g^{ij}y_{,i}y_{,j} = -2\int I(y)dy \equiv -2J(y). \tag{2.10}$$

If V_n is an E_n or S_2, $a=$const. For this case, it follows from (2.9) and (2.10) that

$$I(y) = ay + f, \qquad g^{ij}y_{,i}y_{,j} = -(ay^2 + 2fy + \bar{a}), \tag{2.11}$$

where f and $\bar{a}$ are two constants. Summing up the preceding results we have

THEOREM 2.3. *Let a V_n admit a solution y of* (2.1). *Then equations* (2.9) *and*

(2.10) *hold if* V_n *is an* V_2; *and equations* (2.9), (2.10) *and* (2.11) *hold if* V_n *is an* E_n $(n>2)$ *or* S_2.

To find a meaning of the constant $\bar{a}$ appearing in (2.11), we use (2.1), (2.2) and $(2.11)_2$ in (2.3), and obtain

$$\bar{R}_{ij} = -(n-1)\bar{a}y^{-2}g_{ij}. \tag{2.12}$$

This shows that $\bar{g}_{ij}=y^{-2}g_{ij}$ is the fundamental tensor of an $E_n(\bar{a})$ or $S_2(\bar{a})$. Hence we have

THEOREM 2.4. *If* g_{ij} *is the fundamental tensor of an* $E_n(a)$ *or* $S_2(a)$ *and* y *is a solution of* (2.1), *then* $y^{-2}g_{ij}$ *is the fundamental tensor of an* $E_n(\bar{a})$ *or* $S_2(\bar{a})$, *where* $\bar{a}$ *is determined from* $(2.11)_2$.

We can also prove that

THEOREM 2.5. *If* $g_{ij}=y^2\bar{g}_{ij}$ *is the fundamental tensor of an* S_2 *and* y *satisfies the equation*

$$\left(\frac{1}{y}\right)_{/ij} = -\bar{l}\bar{g}_{ij},$$

where the solidus denotes covariant differentiation with respect to $\bar{g}_{ij}$ *and* $\bar{l}$ *is an unspecified scalar, then* $\bar{g}_{ij}$ *is also the fundamental tensor of an* S_2.

Proof. On account of the preceding theorem, we need show only that (2.1) is satisfied. Now it can be easily verified that if w_i is any vector, then its covariant derivatives $w_{i/j}$ and $w_{i,j}$ taken with respect to $\bar{g}_{ij}$ and $g_{ij}=y^2\bar{g}_{ij}$, respectively, are related by

$$w_{i/j} = w_{i,j} + (w_{,i}y_{,j} + w_{,j}y_{,i})y^{-1} - \bar{g}_{ij}\bar{g}^{hk}y_{,h}y_{,k}. \tag{2.13}$$

Therefore we have, from hypothesis and by (2.13),

$$\begin{aligned} -\bar{l}\bar{g}_{ij} = \left(\frac{1}{y}\right)_{/ij} &= -\frac{y_{/ij}}{y^2} + \frac{2y_{/i}y_{/j}}{y^3} \\ &= -\frac{1}{y^2}\left(y_{,ij} + \frac{2y_{,i}y_{,j}}{y} - \bar{g}_{ij}\bar{g}^{hk}y_{,h}y_{,k}\right) + \frac{2y_{,i}y_{,j}}{y^3} \\ &= -\frac{y_{,ij}}{y^2} + \bar{g}_{ij}\frac{\bar{g}^{hk}y_{,h}y_{,k}}{y^2}, \end{aligned}$$

which shows that y satisfies an equation of the form (2.1), as was to be proved.

Canonical forms for the fundamental tensor of an E_n which admits a solution y of equation (2.1) have been given by Brinkmann [3]. We shall not enter into the detail of his results, but merely mention the main fact that the construction of such a canonical form depends, according as $g^{ij}y_{,i}y_{,j}\neq 0$ or $=0$, on the fundamental tensor of an arbitrary E_{n-1} or on the fundamental

tensor of an E_{n-2} which contains a parameter and satisfies certain differential equations.

If a V_n with nonzero scalar curvature a, which may or may not be constant, admits a solution y of (2.1), then it follows from (2.9) and (2.10) that $g^{ij}y_{,i}y_{,j}\neq 0$. Consequently, we can show, by following Brinkmann's method, that the fundamental form of V_2 can be reduced to

$$(2.14)\qquad -\frac{(dx^1)^2}{2J(x^1)}+2eJ(x^1)(dx^2)^2,$$

where $e=\pm 1$ and $J(x^1)$ is defined by (2.10). Conversely, if the fundamental form of a V_2 is of the form (2.14), where $J(x^1)$ is any function of x^1, then $y=x^1$ is a solution of (2.1).

Finally, we remark that for any V_n the following particular case of equation (2.1):

$$(2.15)\qquad y_{,ij}=0$$

admits a solution, if and only if V_n has a family of parallel totally geodesic hypersurfaces, or what amounts to the same thing, a field of parallel vectors.

3. **Fundamental relations.** From (1.1), it is evident that the components of the conformally separable tensor $^*g_{\alpha\beta}$ and those of its component tensors $^*g_{ij}$ and $^*g_{pq}$ are connected by

$$(3.1)\qquad \begin{array}{lll} {}^*g_{ip}={}^*g^{ip}=0, & {}^*g^{ij}\,{}^*g_{ik}=\delta^j_k, & {}^*g^{pq}\,{}^*g_{pr}=\delta^q_r,\\ {}^*g_{ij}=\rho^{-2}g_{ij}, & {}^*g^{ij}=\rho^2 g^{ij}, & g^{ij}g_{ik}=\delta^j_k,\\ {}^*g_{pq}=\sigma^{-2}g_{pq}, & {}^*g^{pq}=\sigma^2 g^{pq}, & g^{pq}g_{pr}=\delta^q_r, \end{array}$$

where δ^j_k and δ^q_r are Kronecker deltas. Denoting by $^*\Gamma^\gamma_{\alpha\beta}$, $'\Gamma^k_{ij}$, $'\Gamma^r_{pq}$ the Christoffel symbols of the second kind for the fundamental tensors $^*g_{\alpha\beta}$, $^*g_{ij}$, $^*g_{pq}$, respectively, we have [14, (3.3)]

$$(3.2)\qquad \begin{array}{lll} {}^*\Gamma^k_{ij}={}'\Gamma^k_{ij}, & {}^*\Gamma^p_{ij}={}^*g_{ij}\rho^p, & {}^*\Gamma^k_{pj}=-\rho_p\delta^k_j,\\ {}^*\Gamma^r_{pq}={}'\Gamma^r_{pq}, & {}^*\Gamma^h_{pq}={}^*g_{pq}\sigma^h, & {}^*\Gamma^r_{iq}=-\sigma_i\delta^r_q, \end{array}$$

where

$$(3.3)\qquad \begin{array}{ll} \rho_p=\partial_p\log\rho, & \sigma_i=\partial_i\log\sigma,\\ \rho^p={}^*g^{pq}\rho_q, & \sigma^i={}^*g^{ij}\sigma_j. \end{array}$$

The Riemann tensor of $^*g_{\alpha\beta}$ is defined by

$${}^*R^\epsilon_{\alpha\beta\gamma}=\partial_\beta{}^*\Gamma^\epsilon_{\alpha\gamma}-\partial_\gamma{}^*\Gamma^\epsilon_{\alpha\beta}+{}^*\Gamma^\delta_{\alpha\gamma}{}^*\Gamma^\epsilon_{\delta\beta}-{}^*\Gamma^\delta_{\alpha\beta}{}^*\Gamma^\epsilon_{\delta\gamma}.$$

If the values of $^*\Gamma^\gamma_{\alpha\beta}$ given by (3.2) are used, then with some calculation we find

$$
\begin{aligned}
&{}^*R^l_{ijk} = {}'R^l_{ijk} - \rho^p\rho_p(\delta^l_j{}^*g_{ik} - \delta^l_k{}^*g_{ij}),\\
&{}^*R^l_{ijp} = (\delta^l_j\delta^k_i - {}^*g^{lk}{}^*g_{ij})(\partial_k\rho_p + \rho_p\sigma_k) = - {}^*R^l_{ipj},\\
&{}^*R^l_{pjk} = \delta^l_j(\partial_k\rho_p + \rho_p\sigma_k) - \delta^l_k(\partial_j\rho_p + \rho_p\sigma_j),\\
(3.4)\qquad &{}^*R^l_{ipq} = 0,\\
&{}^*R^l_{pqk} = \delta^l_k\rho\left(\frac{1}{\rho}\right)_{;pq} + {}^*g_{pq}{}^*g^{lj}\sigma\left(\frac{1}{\sigma}\right)_{;kj} = - {}^*R^l_{pkq},\\
&{}^*R^l_{pqr} = {}^*g^{hl}[{}^*g_{rp}(\partial_q\sigma_h + \rho_q\sigma_h) - {}^*g_{pq}(\partial_r\sigma_h + \rho_r\sigma_h)],
\end{aligned}
$$

where, as well as in what follows, the semi-colon denotes covariant differentiation with respect to $^*g_{ij}$ or $^*g_{pq}$, and $'R^l_{ijk}$ is the Riemann tensor of $^*g_{ij}$. In deriving (3.4) use has been made of the following formulae:

$$
\begin{aligned}
&{}'\Gamma^l_{ij} = \Gamma^l_{ij} - \delta^l_i\rho_j - \delta^l_j\rho_i + g_{ij}g^{lk}\rho_k \qquad [5,\ \text{p. } 89\ (28.3)],\\
(3.5)\qquad &\partial_k({}^*g_{pq}\sigma^l) = {}^*g_{pq}(-2\sigma_k\sigma^l - {}^*g^{li}{}'\Gamma^j_{ki}\sigma_j - {}'\Gamma^l_{ki}\sigma^i + {}^*g^{lj}\partial_k\sigma_j),\\
&\partial_q\rho_p - {}'\Gamma^r_{pq}\rho_r - \rho_p\rho_q = -\rho\left(\frac{1}{\rho}\right)_{;pq},
\end{aligned}
$$

where Γ^l_{ij} denotes the Christoffel symbols of the second kind for g_{ij}. The expressions for the remaining components of $^*R^\epsilon_{\alpha\beta\gamma}$ are obtained from (3.4) by interchanging the two sets of indices (i, j, k, l) and (p, q, r, s). We remark that (3.4) can be shown to be identical with the Gauss-Codazzi-Ricci equations [5, pp. 162–163, (47.11), (47.12), (47.14)] for the subspaces x^r=const. in the V_m with fundamental tensor $^*g_{\alpha\beta}$.

The components of the Ricci tensor $^*R_{\alpha\beta} = {}^*R^\gamma_{\alpha\beta\gamma}$ of $^*g_{\alpha\beta}$ are readily found from (3.4) by contraction; they are

$$
\begin{aligned}
&-{}^*R_{ip} = (n-1)\partial_i\rho_p + (m-n-1)\partial_p\sigma_i + (m-2)\rho_p\sigma_i,\\
&{}^*R_{ij} = {}'R_{ij} + (m-n)\sigma\left(\frac{1}{\sigma}\right)_{;ij}\\
(3.6)\qquad &\qquad + g_{ij}g^{pq}\left[\rho\left(\frac{1}{\rho}\right)_{;pq} + (n-1)\rho^2\left(\frac{1}{\rho}\right)_{;p}\left(\frac{1}{\rho}\right)_{;q}\right],\\
&{}^*R_{pq} = {}'R_{pq} + n\rho\left(\frac{1}{\rho}\right)_{;pq}\\
&\qquad + g_{pq}g^{ij}\left[\sigma\left(\frac{1}{\sigma}\right)_{;ij} + (m-n-1)\sigma^2\left(\frac{1}{\sigma}\right)_{;i}\left(\frac{1}{\sigma}\right)_{;j}\right],
\end{aligned}
$$

where $'R_{ij}$ ($'R_{pq}$) is zero or is the Ricci tensor of $^*g_{ij}$ ($^*g_{pq}$) according as $^*g_{ij}$ ($^*g_{pq}$) is of dimension 1 or greater than 1.

II. CONFORMALLY SEPARABLE TENSOR OF THE TYPE ($n>1$, $m-n=1$)[11]

4.1. Scalar curvatures of totally umbilical hypersurfaces in an E_m. A conformally separable tensor of the type ($n>1$, $m-n=1$) may be taken as

$$(4.1)\qquad {}^*g_{\alpha\beta}=\begin{bmatrix}\rho^{-2}g_{ij} & 0\\ 0 & e\sigma^{-2}\end{bmatrix},\quad e=\pm 1,\qquad \begin{matrix}\alpha,\beta,\gamma,\delta=1,\cdots,m\ (=n+1),\\ i,j,k,l=1,\cdots,n,\end{matrix}$$

for, we may suppose that $\pm g_{mm}$ has been absorbed in σ^{-2}. From (3.6) it follows that the condition

$$(4.2)\qquad {}^*R_{\alpha\beta}=-(m-1)c{}^*g_{\alpha\beta},\qquad c=\text{const.},$$

for the tensor (4.1) to represent an $E_m(c)$ is

$$(4.3)\qquad \begin{aligned}&\partial_i\rho_m+\rho_m\sigma_i=0,\\ &'R_{ij}+\sigma\left(\frac{1}{\sigma}\right)_{;ij}+{}^*g_{ij}{}^*g^{mm}\left[\rho\left(\frac{1}{\rho}\right)_{;mm}+(m-2)\rho^2\left(\frac{1}{\rho}\right)_{;m}\left(\frac{1}{\rho}\right)_{;m}\right]\\ &\qquad\qquad=-(m-1)c{}^*g_{ij},\\ &(m-1)\rho\left(\frac{1}{\rho}\right)_{;mm}+{}^*g_{mm}{}^*g^{ij}\sigma\left(\frac{1}{\sigma}\right)_{;ij}=-(m-1)c{}^*g_{mm},\end{aligned}$$

where, we repeat, ${}^*g^{mm}=1/{}^*g_{mm}=e\sigma^2$, $'R_{ij}$ is the Ricci tensor of ${}^*g_{ij}$, and the semi-colon denotes covariant differentiation with respect to ${}^*g_{ij}$ or ${}^*g_{mm}$.

If we write

$$(4.4)\qquad {}^*g^{ij}\,'R_{ij}=-(m-1)(m-2){}^*a(x^k,x^m),$$

then, by definition, ${}^*a(x^k,x^m)$ is the scalar curvature of ${}^*g_{ij}=\rho^{-2}g_{ij}$. We shall now prove that ${}^*a(x^k,x^m)$ is independent of x^k.

Transvecting $(4.3)_2$ and $(4.3)_3$ by ${}^*g^{ij}$ and ${}^*g^{mm}$ respectively, and taking account of (4.4), we obtain

$$(4.5)\qquad \begin{aligned}&-(m-1)(m-2){}^*a(x^k,x^m)+{}^*g^{ij}\sigma\left(\frac{1}{\sigma}\right)_{;ij}\\ &\qquad+(m-1){}^*g^{mm}\left[\rho\left(\frac{1}{\rho}\right)_{;mm}+(m-2)\rho^2\left(\frac{1}{\rho}\right)_{;m}\left(\frac{1}{\rho}\right)_{;m}\right]\\ &\qquad\qquad=-(m-1)^2c,\\ &(m-1){}^*g^{mm}\rho\left(\frac{1}{\rho}\right)_{;mm}+{}^*g^{ij}\sigma\left(\frac{1}{\sigma}\right)_{;ij}=-(m-1)c.\end{aligned}$$

When the latter equation is subtracted from the former, and $e\sigma^2$ is used in place of ${}^*g^{mm}$, the result is

[11] In §§4.1 and 4.2 we do not confine ourselves to properly conformally separable tensors, but a complete discussion of improperly conformally separable tensors is reserved for §§7 and 8.

$$ {}^*a(x^k, x^m) - e(\sigma\rho_m)^2 = c. \tag{4.6} $$

If $\partial_m\rho = 0$, this reduces to

$$ {}^*a(x^k, x^m) = c. \tag{4.6'} $$

If $\partial_m\rho \neq 0$, then equation $(4.3)_1$ can be written

$$ \frac{\partial_i\rho_m}{\rho_m} + \sigma_i = 0, \quad \text{that is,} \quad \partial_i \log(\sigma\rho_m) = 0, $$

which gives

$$ \sigma\rho_m = z(x^m), \tag{4.7} $$

where z is a function of x^m alone. Therefore equation (4.6) becomes

$$ {}^*a(x^k, x^m) - e[z(x^m)]^2 = c. \tag{4.8} $$

This equation and (4.6′) show that ${}^*a(x^h, x^m)$ does not depend on x^k, as was to be proved. Hence

THEOREM 4.1. *If a conformally separable tensor of the type $(n>1, m-n=1)$ represents an E_m, then its first component tensor represents V_n's of constant scalar curvatures.*

And geometrically[12],

THEOREM 4.1′. *If in an E_m there exists a one-parameter family of totally umbilical hypersurfaces, then these hypersurfaces are conformal to one another and each of them has constant scalar curvature. If, in particular, the family consists of totally geodesic hypersurfaces, then they are isometric to one another and their constant scalar curvatures are all equal to the scalar curvature of E_m.*

The latter part of this theorem follows from (4.6′).

4.2. **An imbedding theorem.** Continuing our discussion, we now write ${}^*a(x^m) = {}^*a(x^k, x^m)$, and proceed to prove the following

THEOREM 4.2. *In order that the conformally separable tensor* (4.1) *with $\partial_m\rho \neq 0$ may represent an $E_m(c)$ and its first component tensor ${}^*g_{ij} = \rho^{-2}g_{ij}$ E_n's or S_2's of scalar curvatures ${}^*a(x^m)$, it is necessary and sufficient that when*

$$ \rho(x^k, x_0^m) = 1 \tag{4.9} $$

is assumed, the equations

---

(12) A more general result can be obtained by using the Gauss-Codazzi equations of a V_n in an E_{n+1}. Indeed we can prove that *if an E_{n+1} has a totally umbilical hypersurface V_n, then the scalar curvature of V_n is constant.* But we shall not go farther with this result, because Theorem 4.1′ serves only as a preliminary to the imbedding Theorem 4.2′.

$$R_{ij} = -(n-1)ag_{ij}, \qquad a = \text{const.}, \tag{4.10}$$

$$\rho_{,ij} \sim g_{ij}, \tag{4.11}$$

$$e\sigma^{-2} = \frac{(\rho_m)^2}{{}^*a(x^m) - c} \tag{4.12}$$

be satisfied.

Here we write (4.11) to mean the equation $\rho_{,ij} = -\omega g_{ij}$, where ω is an unspecified scalar. This notation will be used whenever desirable; it enables us to avoid the unnecessary introduction of many symbols to represent scalar factors of proportionality.

We know that in a V_m the hypersurfaces $x^m = \text{const.}$, whose (first) fundamental tensors are nonsingular, are totally umbilical if and only if the fundamental tensor of V_m can be reduced to the form (4.1) [5, pp. 144, 182]. We also know from Theorem 2.1 that when $n > 2$ and (4.9) is satisfied, equations (4.10), (4.11) are the conditions that

$$'R_{ij} = -(n-1){}^*a(x^m){}^*g_{ij}, \tag{4.13}$$

that is that ${}^*g_{ij} = \rho^{-2}g_{ij}$ represents E_n's. Accordingly Theorem 4.2 may be stated geometrically as follows.

THEOREM 4.2'. *Let c be any constant and let ${}^*g_{ij} = [\rho(x^k, x^m)]^{-2}g_{ij}(x^k)$ with x^m as parameter and $\partial_m\rho \neq 0$ represent ∞^1 V_n's whose scalar curvatures are not all equal to c. If the V_n's are E_n's $(n > 2)$, there exists an $E_{n+1}(c)$ in which they are imbedded isometrically as totally umbilical hypersurfaces. If the V_n's are S_2's, a necessary and sufficient condition that they may be imbedded isometrically in an $S_3(c)$ as totally umbilical surfaces is that, when $\rho(x^k, x_0^m) = 1$ is assumed, the tensor $\rho_{,ij}$ differ from g_{ij} by a scalar factor. The fundamental tensor of the imbedding $E_{n+1}(c)$ or $S_3(c)$, if it exists, is*

$${}^*g_{\alpha\beta} = \begin{bmatrix} \rho^{-2}g_{ij} & 0 \\ 0 & (\partial_n \log \rho)^2/[{}^*a(x^m) - c] \end{bmatrix},$$

where ${}^*a(x^m)$ denotes the scalar curvatures of the given E_n's or S_2's.

We shall now prove Theorem 4.2. Since by hypothesis $\partial_m\rho \neq 0$, equations (4.6), (4.7) and (4.8) are consequences of (4.3), as we have seen in §4.1. If we solve (4.6) for $e\sigma^{-2}$, the result is (4.12). Hence Theorem 4.2 will be proved if we can show that in consequence of[12] (4.7), (4.8), (4.9) and (4.13), equations (4.3) reduce to (4.11).

Now equation $(4.3)_1$ is satisfied because of (4.7). When (4.13) is used, equation $(4.3)_2$ becomes

---

(12) We observe that equations (4.6)–(4.8) and (4.12) are unaffected by the supposition (4.9).

$$\sigma\left(\frac{1}{\sigma}\right)_{;ij} = \left\{ -(m-1)c + (n-1)^*a(x^m) - {}^*g^{mm}\left[\rho\left(\frac{1}{\rho}\right)_{;mm} + (m-2)\rho^2\left(\frac{1}{\rho}\right)_{;m}\left(\frac{1}{\rho}\right)_{;m}\right]\right\}{}^*g_{ij},$$

which is evidently equivalent to $(4.5)_1$ and

$$(4.14)\qquad \left(\frac{1}{\sigma}\right)_{;ij} \sim {}^*g_{ij}.$$

But equation $(4.5)_1$ is a consequence of (4.7), (4.8) and $(4.3)_3$, as is evident from the way in which (4.7) and (4.8) were derived. Thus, because of (4.7), (4.8) and (4.13), equations (4.3) are equivalent to (4.14) and $(4.3)_3$. In what follows we shall reduce the latter two equations successively to (4.16), then to (4.20), and finally to (4.11).

Comparing (4.14) and (4.13) with (2.1) and (2.2), it follows from $(2.11)_1$ that equation (4.14) can be written

$$(4.14')\qquad \left(\frac{1}{\sigma}\right)_{;ij} = -\left[{}^*a(x^m)\frac{1}{\sigma} + w(x^m)\right]{}^*g_{ij},$$

where $w(x^m)$ is a function of x^m alone. In virtue of this, equation $(4.3)_3$ becomes

$$(4.15)\qquad e\sigma^2\rho\left(\frac{1}{\rho}\right)_{;mm} + c - {}^*a(x^m) - \sigma w(x^m) = 0.$$

Now

$$\left(\frac{1}{\rho}\right)_{;m} = -\frac{\rho_m}{\rho}$$

$$\begin{aligned}\left(\frac{1}{\rho}\right)_{;mm} &= \partial_m\left(-\frac{\rho_m}{\rho}\right) - {}'\Gamma^m_{mm}\left(-\frac{\rho_m}{\rho}\right)\\ &= \frac{\rho_m}{\rho}\left(-\partial_m \log\frac{\rho_m}{\rho} + (1/2)\partial_m \log {}^*g_{mm}\right)\\ &= \frac{\rho_m}{\rho}\partial_m \log\frac{\rho}{\sigma\rho_m} = \frac{\rho_m}{\rho}\left[\rho_m - \frac{z'(x^m)}{z(x^m)}\right]\end{aligned}$$

by (4.7), where and in what follows, the prime denotes differentiation. Substituting the above expression in (4.15) and then using (4.7), (4.8) and the equation obtained by differentiating (4.8), we find

$$(4.15')\qquad w(x^m) = -ez'(x^m) = -\frac{{}^*a'(x^m)}{2z(x^m)} = -\frac{{}^*a'(x^m)}{2\sigma\rho_m}.$$

Because of this, equation (4.14′) becomes

$$\left(\frac{1}{\sigma}\right)_{;ij} = -\frac{1}{\sigma}\left[{}^*a(x^m) - \frac{{}^*a'(x^m)}{2\rho_m}\right]{}^*g_{ij}. \tag{4.16}$$

This is an equation which, because of (4.7) and (4.8), is equivalent to (4.14) and $(4.3)_3$.

We now express (4.16) directly in terms of ρ and g_{ij}. To do this we make use of (4.7) and the following formula, which can easily be proved by means of $(3.5)_1$:

$$\left(\frac{1}{\sigma}\right)_{;ij} = \frac{1}{\rho}\left(\frac{\rho}{\sigma}\right)_{,ij} - \frac{\rho_{,ij}}{\sigma} - g^{hk}\left(\frac{1}{\sigma}\right)_{,h}\rho_{,k}g_{ij}, \tag{4.17}$$

where as usual the comma denotes covariant differentiation with respect to g_{ij}. Equation (4.16) then becomes

$$\begin{aligned}[- {}^*a(x^m)\rho_m + (1/2){}^*a'(x^m) + \rho g^{hk}(\rho_m)_{,h}\rho_{,k}]g_{ij}& \\ = \rho(\partial_m\rho)_{,ij} - (\partial_m\rho)\rho_{,ij} &= \rho^2\partial_m\left(\frac{\rho_{,ij}}{\rho}\right),\end{aligned} \tag{4.18}$$

because $(\partial_m\rho)_{,ij} = \partial_m(\rho_{,ij})$. On account of (4.9), the scalar curvatures $a = {}^*a(x_0^m)$ and ${}^*a(x^m)$ of g_{ij} and $\rho^{-2}g_{ij}$ are connected by [5, p. 90, (28.7)]

$${}^*a(x^m) = a\rho^2 + \frac{2}{n}\rho g^{hk}\rho_{,hk} - g^{hk}\rho_{,h}\rho_{,k}. \tag{4.19}$$

By use of this equation and its partial derivative with respect to x^m, we can easily verify that the coefficient of g_{ij} in (4.18) is

$$\frac{1}{n}g^{hk}[\rho(\partial_m\rho)_{,hk} - (\partial_m\rho)\rho_{,hk}].$$

This shows that, when (4.9) is supposed, equation (4.18), and hence also (4.16), are equivalent to

$$\partial_m\left(\frac{\rho_{,ij}}{\rho}\right) \sim g_{ij}. \tag{4.20}$$

Finally, to reduce this to (4.11), we integrate it with respect to x^m and obtain

$$\frac{\rho_{,ij}}{\rho} + T_{ij} \sim g_{ij}, \tag{4.21}$$

where T_{ij} is an integration tensor independent of x^m. Now it follows from the very definition of partial differentiation that for any function $\phi(x^k, x^m)$ of x^k and x^m,

$$[\partial_i\phi(\overset{k}{x},\overset{m}{x})]_{x^m=x_0^m}=\partial_i\phi(\overset{k}{x},\overset{m}{x_0}).$$

Therefore, in consequence of (4.9), we have $(\rho_{,i})_{x^m=x_0^m}=0$, $(\rho_{,ij})_{x^m=x_0^m}=0$, and whence, if we put $x^m=x_0^m$ in (4.21), the result is $T_{ij}\sim g_{ij}$. This shows that (4.21) and hence also (4.20) are equivalent to (4.11). The proof of our theorem has thus been completed.

Added in proof. In connection with Theorem 4.2′ I may mention that in a forthcoming paper of mine [17] a necessary and sufficient condition is obtained for a V_n to be imbeddable in an E_{n+1} as a member of ∞^1 totally umbilical hypersurfaces. There Theorem 4.2′ appears as a corollary to a more general result, and all the conformal-Euclidean V_n's which satisfy this condition of imbeddability are determined.

III. Properly conformally separable tensors of the type $(n>1,\ m-n>1)$

5.1. **An auxiliary theorem.** In this section we shall consider the conformally separable tensor

$$\text{(5.1)}\qquad {}^*g_{\alpha\beta}=\begin{bmatrix}\rho^{-2}g_{ij} & 0\\ 0 & \sigma^{-2}g_{pq}\end{bmatrix},\qquad \begin{aligned}\alpha,\beta,\gamma,\delta&=1,\cdots,m,\\ i,j,k,l&=1,\cdots,n,\\ p,q,r,s&=n+1,\cdots,m,\end{aligned}$$

where

$$g_{ij}=g_{ij}(x^k),\quad g_{pq}=g_{pq}(x^r);\qquad \rho=\rho(x^\alpha),\quad \sigma=\sigma(x^\alpha),\quad \partial_p\rho\neq 0,\quad \partial_i\sigma\neq 0.$$

For (5.1), we have (cf. (3.6))

$$\text{(5.2)}\qquad \begin{aligned}-{}^*R_{ip}&=(n-1)\partial_i\rho_p+(m-n-1)\partial_p\sigma_i+(m-2)\rho_p\sigma_i,\\ {}^*R_{ij}&={}'R_{ij}+(m-n)\sigma\left(\frac{1}{\sigma}\right)_{;ij}\\ &\quad+{}^*g_{ij}{}^*g^{pq}\left[\rho\left(\frac{1}{\rho}\right)_{;pq}+(n-1)\rho^2\left(\frac{1}{\rho}\right)_{;p}\left(\frac{1}{\rho}\right)_{;q}\right],\\ {}^*R_{pq}&={}'R_{pq}+n\rho\left(\frac{1}{\rho}\right)_{;pq}\\ &\quad+{}^*g_{pq}{}^*g^{ij}\left[\sigma\left(\frac{1}{\sigma}\right)_{;ij}+(m-n-1)\sigma^2\left(\frac{1}{\sigma}\right)_{;i}\left(\frac{1}{\sigma}\right)_{;j}\right],\end{aligned}$$

where the signs (;) (′) indicate, respectively, the covariant differentiation and the Ricci tensor referred to ${}^*g_{ij}$ or ${}^*g_{pq}$. We suppose as usual that

$$\text{(5.3)}\qquad \rho(\overset{k}{x},\overset{r}{x_0})=1,\qquad \sigma(\overset{k}{x_0},\overset{r}{x})=1.$$

To establish our main result, Theorem 5.2, in the latter part of this section, we need the following auxiliary

THEOREM 5.1. *For the properly conformally separable tensor* (5.1) *of the type* ($n>1$, $m-n>1$) *with* (5.3) *satisfied*:

(i) *The system of equations*

$$-{}^*R_{ip} = (n-1)\partial_i\rho_p + (m-n-1)\partial_p\sigma_i + (m-2)\rho_p\sigma_i = 0 \tag{5.4}$$

is equivalent to

$$\begin{gathered}\rho = \rho(y, z), \qquad \sigma = \sigma(y, z),\\ \rho(y, z_0) = 1, \qquad \sigma(y_0, z) = 1,\\ (n-1)\partial_y\rho_z + (m-n-1)\partial_z\sigma_y + (m-2)\rho_z\sigma_y = 0\text{[14]},\end{gathered} \tag{5.4'}$$

where $y=y(x^k)$, $z=z(x^r)$ are any functions of the arguments indicated, and $y_0=y(x_0^k)$, $z_0=z(x_0^r)$.

(ii) *If the tensors g_{ij} and g_{pq} are considered as given, the following system of equations in the unknown functions $\rho(x^\alpha)$ and $\sigma(x^\alpha)$*:

$$^*R_{ip} = 0, \tag{5.5$_1$}$$

$$^*R_{ij} - {}'R_{ij} \sim {}^*g_{ij}, \qquad {}^*R_{pq} - {}'R_{pq} \sim {}^*g_{pq} \tag{5.5$_2$}$$

is equivalent to the system of equations consisting of (5.4′) and

$$\begin{gathered}\left(\frac{\rho}{\sigma}\right)^2\partial_y\sigma = Q(z), \qquad \left(\frac{\sigma}{\rho}\right)^2\partial_z\rho = J(y),\\ y_{,ij} \sim g_{ij}, \qquad z_{,pq} \sim g_{pq},\end{gathered} \tag{5.5'}$$

in the unknown functions $y=y(x^k)$, $z=z(x^r)$, $\rho=\rho(y, z)$, $\sigma=\sigma(y, z)$, where J and Q are any functions of the arguments indicated.

5.2. **Proof of Theorem** 5.1 (i). Equation (5.4) can be written

$$-(m-2)\rho_p\sigma_i = \partial_p\partial_i[(n-1)\log\rho + (m-n-1)\log\sigma].$$

Differentiating this partially with respect to x^q and then taking the alternation in the indices p and q, we get[15]

$$-(m-2)(\sigma_i\partial_{[q}\rho_{p]} + \rho_{[p}\partial_{q]}\sigma_i) = 0,$$

that is,

$$\rho_{[p}\partial_{q]}\sigma_i = 0. \tag{5.6$_a$}$$

When the value of $\partial_q\sigma_i$ from (5.4) is substituted (which is possible because $m-n>1$), this becomes

[14] Here we write $\partial_y = \partial/\partial y$, $\partial_z = \partial/\partial z$, $\rho_z = \partial_z \log \rho$, $\sigma_y = \partial_y \log \sigma$.

[15] We write, for example, $\partial_{[q}\rho_{p]} = \partial_q\rho_p - \partial_p\rho_q$.

$$-\rho_{[p}[(n-1)\partial_{q]}\rho_i + (m-2)\rho_{q]}\sigma_i] = 0,$$

which, since $n>1$, reduces to

$$\rho_{[p}\partial_{q]}\rho_i = 0.$$

This shows that a function $\theta(x^k, x^r)$ exists such that

$$\partial_i\rho_p = \partial_p\rho_i = (\partial_i \log \theta)\rho_p,$$

and hence

$$\rho_p = \theta(x^k, x^r)w_p(x^r),$$

where $w_p(x^r)$ are $m-n$ functions of x^r alone. Now from the very definition of partial derivative, we have that for any fixed values x_1^a of x^a,

$$\partial_p \log \rho(\overset{k}{x_1}, \overset{r}{x}) = [\partial_p \log \rho(\overset{k}{x}, \overset{r}{x})]_{x^k=x_1^k}.$$

Therefore, if we write $z\equiv\rho(x_1^k, x^r)$ and remember that $\rho_p=\partial_p \log \rho$, then it follows from the two preceding equations that

$$\partial_p \log z = \theta(\overset{k}{x_1}, \overset{r}{x})w_p(\overset{r}{x}) = \frac{\theta(\overset{k}{x_1}, \overset{r}{x})}{\theta(x^k, x^r)}\rho_p,$$

that is,

$$\partial_p \log \rho = \frac{\theta(x^k, x^r)}{\theta(x_1^k, x^r)}\partial_p \log z,$$

which shows that ρ can be expressed in terms of x^k and z alone; thus,

$$(5.7)_a \qquad \rho = \rho(x^k, z).$$

Since (5.4) as well as the hypothesis following (5.1) remain the same when $\rho, \sigma; n, m-n; i, p$ are interchanged, we have, by symmetry,

$$(5.7)_b \qquad \sigma = \sigma(y, x^r),$$

where y is defined by $y\equiv\sigma(x^k, x_1^r)$.

From (5.7) we have

$$\rho_p = \rho_z\partial_p z, \quad \partial_i\rho_p = (\partial_i\rho_z)\partial_p z; \qquad \sigma_i = \sigma_y\partial_i y, \quad \partial_p\sigma_i = (\partial_p\sigma_y)\partial_i y,$$

where $\rho_z=\partial_z \log \rho$, $\sigma_y=\partial_y \log \sigma$. Using these in

$$(5.6)_b \qquad \sigma_{[i}\partial_{j]}\rho_p = 0,$$

which is the symmetric expression of $(5.6)_a$, we find

$$\sigma_y(\partial_p z)(\partial_{[i}y)\partial_{j]}\rho_z = 0.$$

Since by hypothesis $\partial_p\rho\neq 0$, $\partial_i\sigma\neq 0$, so that $\sigma_y\partial_p z\neq 0$, the above equations are equivalent to

$$(\partial_{[i}y)\partial_{j]}\rho_z = 0.$$

From this it follows that ρ_z can be expressed in terms of x^r and y alone. But on the other hand, $(5.7)_a$ shows that ρ_z can be expressed in terms of x^k and z alone. Therefore ρ_z is a function of y and z alone, and hence ρ must be of the form

$$\rho = F(x^k)\phi(y, z).$$

Taking (5.3) into account, we have

$$\rho = \frac{\rho(x^k, x^r)}{\rho(x^k, x^r_0)} = \frac{F(x^k)\phi(y, z)}{F(x^k)\phi(y, z_0)},$$

where $z_0 = z(x^r_0)$. Thus

$$(5.8)_a \qquad \rho = \rho(y, z), \qquad \rho(y, z_0) = 1$$

and by symmetry,

$$(5.8)_b \qquad \sigma = \sigma(y, z), \qquad \sigma(y_0, z) = 1.$$

Now for ρ and σ of the form (5.8), equations (5.4) become, after omitting the non-vanishing factor $(\partial_i y)(\partial_p z)$,

$$(5.9) \qquad (n - 1)\partial_y\rho_z + (m - n - 1)\partial_z\sigma_y + (m - 2)\rho_z\sigma_y = 0.$$

Equations (5.8) and (5.9) are identical with (5.4′), and therefore Theorem 5.1 (i) is proved.

Remark. For any V_m with fundamental tensor (5.1), $^*R_{ip} = 0$ is the condition that there be n independent congruences of Ricci curves of V_m lying in the subspaces $x^r = \text{const}$. Hence from (5.8) we have incidentally:

If, in a V_m with a properly conformally separable fundamental tensor of the type $(n > 1,\ m - n > 1)$, the subspaces $x^r = \text{const.}$ contain n independent congruences of Ricci curves of V_m, then the ∞^{m-n} subspaces $x^r = \text{const.}$ and the ∞^n subspaces $x^k = \text{const.}$ consist of ∞^1 families of ∞^{m-n-1} isometric V_n's and ∞^1 families of ∞^{n-1} isometric V_{m-n}'s, respectively.

5.3. **Proof of Theorem** 5.1 (ii). By Theorem 5.1 (i), equation $(5.5)_1$, which is identical with (5.4), is equivalent to (5.4′). If

$$y = y('y), \qquad z = z('z)$$

is any nonsingular transformation from y, z to $'y, 'z$, then (5.4′) become

$$\rho = \rho[y('y), z('z)], \qquad \sigma = \sigma[y('y), z('z)],$$

$$(5.4'') \qquad \rho[y('y), z('z_0)] = 1, \qquad \sigma[y('y_0), z('z)] = 1,$$

$$(n - 1)\partial_{'y}\rho_{'z} + (m - n - 1)\partial_{'z}\sigma_{'y} + (m - 2)\rho_{'z}\sigma_{'y} = 0,$$

where $'y_0$, $'z_0$ are any roots of the equations $y('y) = y_0$, $z('z) = z_0$. Hence

Theorem 5.1(ii) is true if a suitable nonsingular transformation $y=y('y)$, $z=z('z)$ exists such that equations $(5.5)_2$ reduce to

$$(5.5'') \qquad \left(\frac{\rho}{\sigma}\right)^2 \partial_{'y}\sigma = 'Q('z), \qquad \left(\frac{\sigma}{\rho}\right)^2 \partial_{'z}\rho = 'J('y),$$
$$'y_{,ij} \sim g_{ij}, \qquad 'z_{,pq} \sim g_{pq}.$$

We now proceed to prove that this is the case.

By $(5.2)_2$, equation $(5.5)_{2a}$ is equivalent to

$$(5.5''')_{2a} \qquad \left(\frac{1}{\sigma}\right)_{;ij} \sim g_{ij}.$$

When the covariant derivative $(1/\sigma)_{;ij}$ taken with respect to $^*g_{ij}=\rho^{-2}g_{ij}$ is expressed in terms of the covariant derivative $(1/\sigma)_{,ij}$ taken with respect to g_{ij}, equation $(5.5''')_{2a}$ becomes (cf. (2.13))

$$(5.5^{iv})_{2a} \qquad \left(\frac{1}{\sigma}\right)_{,ij} + \rho_i\left(\frac{1}{\sigma}\right)_{,j} + \rho_j\left(\frac{1}{\sigma}\right)_{,i} \sim g_{ij}.$$

By $(5.4')_1$, ρ and σ are functions of y and z, and therefore

$$\left(\frac{1}{\sigma}\right)_{,i} = -\frac{\partial_y\sigma}{\sigma^2}\, y_{,i},$$
$$\left(\frac{1}{\sigma}\right)_{,ij} = -\frac{\partial_y\sigma}{\sigma^2}\, y_{,ij} + 2\frac{(\partial_y\sigma)^2}{\sigma^3}\, y_{,i}y_{,j} - \frac{\partial_y\partial_y\sigma}{\sigma^2}\, y_{,i}y_{,j},$$
$$\rho_i = \frac{\partial_y\rho}{\rho}\, y_{,i}.$$

When these are used in $(5.5^{iv})_{2a}$, the latter becomes

$$-\frac{\partial_y\sigma}{\sigma^2}\, y_{,ij} + \left[2\frac{(\partial_y\sigma)^2}{\sigma^3} - \frac{\partial_y\partial_y\sigma}{\sigma^2} - 2\frac{\partial_y\sigma}{\sigma^2}\frac{\partial_y\rho}{\rho}\right] y_{,i}y_{,j} \sim g_{ij},$$

which, because $\partial_y\sigma \neq 0$ by hypothesis, can be written

$$(5.10) \qquad y_{,ij} + \left(\partial_y \log \frac{\rho^2\partial_y\sigma}{\sigma^2}\right) y_{,i}y_{,j} \sim g_{ij}.$$

A consequence of this equation is

$$(5.11) \qquad \partial_y \log \frac{\rho^2\partial_y\sigma}{\sigma^2} = \text{a function of } y \text{ alone}.$$

For, if we write (5.10) as

$$y_{,ij} + \lambda(y, z)y_{,i}y_{,j} \sim g_{ij},$$

and eliminate $y_{,ij}$ from it and

$$y_{,ij} + \lambda(y, z_1)y_{,i}y_{,j} \sim g_{ij},$$

where z_1 is a constant, we find

$$[\lambda(y, z) - \lambda(y, z_1)]y_{,i}y_{,j} \sim g_{ij}.$$

Since g_{ij} is of rank greater than 1, the coefficient of $y_{,i}\, y_{,j}$ in the above equation must be zero; thus $\lambda(y, z) = \lambda(y, z_1)$, which proves (5.11).

This being the case, we have by integration of (5.11) that

$$(5.12)_a \qquad \frac{\rho^2 \partial_y \sigma}{\sigma^2} = I(y)Q(z).$$

In like manner we derive from $(5.5)_{2b}$ that

$$(5.12)_b \qquad \frac{\sigma^2 \partial_z \rho}{\rho^2} = J(y)P(z).$$

Here I, J, P, Q are some functions of the arguments indicated.

Now consider the functions $'y$, $'z$ introduced (to within integration constants) by

$$(5.13) \qquad 'y = \int I(y)dy, \qquad 'z = \int P(z)dz.$$

Since $\partial_z\rho \neq 0$, $\partial_y\sigma \neq 0$ by hypothesis, it follows from (5.12) that neither $I(y)$ nor $P(z)$ can be identically zero. Consequently, (5.13) define a nonsingular transformation, which evidently carries (5.12) into

$$(5.14)_1 \qquad \frac{\rho^2 \partial_{'y} \sigma}{\sigma^2} = {}'Q('z), \qquad \frac{\sigma^2 \partial_{'z} \rho}{\rho^2} = {}'J('y).$$

Finally, if we recall the way in which (5.10) was derived from $(5.5)_{2a}$, it will at once become obvious that the expression for $(5.5)_{2a}$ in terms of $'y$ and $'z$ is obtained by replacing y by $'y$ in (5.10); that is, (5.13) transforms $(5.5)_{2a}$ into

$$'y_{,ij} + \left(\partial_{'y} \log \frac{\rho^2 \partial_{'y} \sigma}{\sigma^2}\right) 'y_{,i}'y_{,j} \sim g_{ij}.$$

In consequence of $(5.14)_{1a}$, this becomes

$$(5.14)_{2a} \qquad 'y_{,ij} \sim g_{ij}.$$

Similarly, in terms of $'y$ and $'z$, equation $(5.5)_{2b}$ becomes

$$(5.14)_{2b} \qquad 'z_{,pq} \sim g_{pq}.$$

Equations (5.14) are identical with $(5.5'')_2$, which proves our theorem.

Remark. From the above proof, it is easily seen that equations $(5.5)_1$, $(5.5)_{2a}$, by themselves, are equivalent to $(5.4')$, $(5.5')_{1a}$, $(5.5')_{2a}$; and equations $(5.5)_1$, $(5.5)_{2b}$, by themselves, to $(5.4')$, $(5.5')_{1b}$, $(5.5')_{2b}$.

5.4 **An important property.** We are now ready to prove the following

THEOREM 5.2. *If a properly conformally separable tensor* $^*g_{\alpha\beta}$ *of the type* $(n, m-n)$ *represents an* E_m *and each of its component tensors either is of dimension* 2 *or represents* E_n's, *then* $^*g_{\alpha\beta}$ *is conformal to a separable tensor of the type* $(n, m-n)$.

Proof. By supposition, we have

$$(5.15) \qquad {}^*R_{\alpha\beta} = -(m-1)c\, {}^*g_{\alpha\beta}, \qquad c = \text{const.},$$

$$(5.16)_a \qquad {}'R_{ij} = -(n-1){}^*a(x^k, x^r){}^*g_{ij},$$

$$(5.16)_b \qquad {}'R_{pq} = -(m-n-1){}^*b(x^k, x^r){}^*g_{pq}.$$

As a consequence of these equations, equations (5.5) are true. Thus by Theorem 5.1 (ii), two functions $y=y(x^k)$, $z=z(x^r)$ exist such that the following equations are satisfied:

$$(5.17) \qquad \begin{cases} \rho = \rho(y, z), \qquad \sigma = \sigma(y, z), \\ \rho(y, z_0) = 1, \qquad \sigma(y_0, z) = 1, \\ (n-1)\partial_y\rho_z + (m-n-1)\partial_z\sigma_y + (m-2)\rho_z\sigma_y = 0. \end{cases}$$

$$(5.18) \qquad \begin{cases} \left(\dfrac{\rho}{\sigma}\right)^2 \partial_y\sigma = Q(z), \qquad \left(\dfrac{\sigma}{\rho}\right)^2 \partial_z\rho = J(y), \\ y_{,ij} \sim g_{ij}, \qquad z_{,pq} \sim g_{pq}. \end{cases}$$

Our theorem will be proved if we can show that as a result of (5.16), (5.17) and (5.18), the function ρ/σ is of the form $'\rho(y)/'\sigma(z)$. We treat the two cases $n>2$, $m-n>1$ and $n=m-n=2$ separately.

Case 1. $n>2$, $m-n>1$.

Since $n>2$, we have by supposition that $\rho^{-2}g_{ij}$ represents E_n's. Therefore it follows from $(5.16)_a$ and $(5.17)_{1a}$ that $^*a(x^k, x^r) = {}^*a(z)$. On account of $(5.17)_{2a}$, g_{ij} is the fundamental tensor of an E_n, whose scalar curvature is $a={}^*a(z_0)$. Thus, by Theorem 2.1, equation $(5.16)_a$ implies that

$$(5.19) \qquad \rho_{,ij} \sim g_{ij}.$$

In virtue of $(5.17)_{1a}$, this can be written

$$(\partial_y\rho)y_{,ij} + (\partial_y\partial_y\rho)y_{,i}y_{,j} \sim g_{ij},$$

which becomes, because of $(5.18)_{2a}$,

$$(\partial_y\partial_y\rho)y_{,i}y_{,j} \sim g_{ij}.$$

Since g_{ij} is of rank greater than 1, it follows from this that $\partial_y\partial_y\rho=0$, whence

$$\rho = U(z)y + V(z). \tag{5.20}$$

Using this in $(5.18)_{1a}$ we have

$$\partial_y\left(\frac{1}{\sigma}\right) = -Q(Uy+V)^{-2} = \begin{cases} -QV^{-2} & \text{if } U = 0, \\ \dfrac{Q}{U}\,\partial_y(Uy+V)^{-1} & \text{if } U \neq 0. \end{cases}$$

Integrating these with respect to y and then making use of (5.20), we find that for both cases ρ/σ is of the form

$$\frac{\rho}{\sigma} = W(z)y + Z(z). \tag{5.21}$$

The U, V in (5.20) and W, Z in (5.21) are all functions of z alone.

If either W or Z is identically zero, ρ/σ will be of the form $'\rho(y)/'\sigma(z)$, and our theorem is proved.

Now suppose that neither W nor Z is identically zero. When (5.20), (5.21) are used in $(5.17)_2$, the latter reduces to

$$\begin{aligned} &(m-2)(U'Z - V'W)(Wy+Z) \\ &\qquad + (m-n-1)(WZ' - W'Z)(Uy+V) = 0, \end{aligned} \tag{5.22}$$

where the prime denotes differentiation. From this it can be proved that

$$WZ' - W'Z = 0. \tag{5.23}$$

Assume that this is not true. Then since $W\neq 0$, $Z\neq 0$, we have from (5.22) that

$$U/W = V/Z = X(z),$$

where $X\neq 0$, otherwise, $\rho=0$ by (5.20). Now it is easily verified that in consequence of the above equations, (5.22) becomes $(n-1)(WZ'-W'Z)X=0$, which cannot be satisfied. Thus (5.23) is true, and consequently W and Z differ by a constant factor. Hence it follows from (5.21) that ρ/σ is the form $'\rho(y)/'\sigma(z)$, and the proof of Theorem 5.2 for the case $n>2$, $m-n>1$ is completed.

Remark. We observe that in the above proof we made use only of equations $(5.16)_a$, (5.17), $(5.18)_{1a}$, $(5.18)_{2a}$, but not of (5.15), $(5.16)_b$, $(5.18)_{1b}$, $(5.18)_{2b}$. Now by the remark at the end of §5.3, equations (5.17), $(5.18)_{1a}$, $(5.18)_{2a}$ are equivalent to $(5.5)_1$, $(5.5)_{2a}$, namely,

$$^*R_{ip} = 0, \qquad ^*R_{ij} - {}'R_{ij} \sim {}^*g_{ij}.$$

Thus our conclusion that ρ/σ is of the form $'\rho(y)/'\sigma(z)$ is in fact a conse-

quence of the following equations:

$$ {}^*R_{i\nu} = 0, \qquad {}'R_{ij} \sim {}^*g_{ij}, \qquad {}^*R_{ij} \sim {}^*g_{ij}. $$

Hence:

If, at each point P, coordinates x_1^α, of a V_m with properly conformally separable fundamental tensor ${}^*g_{\alpha\beta}$ of the type $(n>2, m-n>1)$, every direction in the $V_n: x^r = x_1^r$ is a Ricci direction both of V_m and of V_n, then ${}^*g_{\alpha\beta}$ is conformal to a separable tensor of the type $(n, m-n)$.

Case 2. $n = m-n = 2$.

In this case, equation (5.19) is in general not true, and we shall base our proof of Theorem 5.2 on (5.17) and (5.18) alone. Equation $(5.17)_3$ is now equivalent to

$$ \partial_y \rho_z + \rho_z \sigma_y = \omega, \qquad \partial_z \sigma_y + \rho_z \sigma_y = -\omega, \tag{5.24} $$

where $\omega = \omega(y, z)$. If $\omega = 0$, then since $\rho_z \neq 0$, $(5.24)_a$ can be written

$$ \partial_y \log (\sigma\rho_z) = 0, $$

which gives us, on integration,

$$ \sigma\rho_z = \frac{1}{{}'\sigma(z)}, \qquad \text{that is,} \qquad \frac{\sigma}{\rho}\,\partial_z\rho = \frac{1}{{}'\sigma(z)}. $$

Comparison of the last equation with $(5.18)_{1b}$ shows that $\sigma/\rho = {}'\sigma(z)J(y)$, which proves our theorem.

We now suppose that $\omega \neq 0$ and always bear in mind that $J = J(y)$, $Q = Q(z)$. Then on account of $(5.18)_1$, equation $(5.24)_a$ can be written

$$ \omega = \frac{J\rho}{\sigma^2}\,\partial_y \log\left(J\frac{\rho}{\sigma}\right) = \frac{J\rho}{\sigma^2}\left(\frac{J'}{J} + \rho_y - \frac{Q\sigma}{\rho^2}\right), $$

that is,

$$ \partial_y\left(\frac{1}{\rho}\right) = \frac{J'}{\rho J} - \frac{Q\sigma}{\rho^3} - \frac{\omega\sigma^2}{J\rho^2}. \tag*{$(5.25)_a$} $$

In like manner, $(5.24)_b$ can be reduced to

$$ \frac{Q\partial_z\sigma}{\rho^2} = -\frac{Q'\sigma}{\rho^2} + \frac{JQ}{\rho\sigma} - \omega. \tag*{$(5.25)_b$} $$

We now find the integrability condition $\partial_z\partial_y(1/\rho) = \partial_y\partial_z(1/\rho)$ for $(5.18)_{1b}$ and $(5.25)_a$. Differentiate $(5.25)_a$ with respect to z and we have

$$ \begin{aligned} \partial_z\partial_y\left(\frac{1}{\rho}\right) = {} & \frac{J'}{J}\partial_z\left(\frac{1}{\rho}\right) - \left[\frac{Q'\sigma}{\rho^3} + \frac{Q\partial_z\sigma}{\rho^3} + \frac{3Q\sigma}{\rho^2}\partial_z\left(\frac{1}{\rho}\right)\right] \\ & - \frac{1}{J}\left[\frac{\sigma^2}{\rho^2}\,\partial_z\omega + \frac{2\omega\sigma\partial_z\sigma}{\rho^2} + \frac{2\omega\sigma^2}{\rho}\partial_z\left(\frac{1}{\rho}\right)\right]. \end{aligned} \tag{5.26} $$

On the other hand we have from $(5.18)_1$

$$(5.27)\quad \partial_x\left(\frac{1}{\rho}\right) = -\frac{J}{\sigma^2}, \qquad \partial_y\partial_x\left(\frac{1}{\rho}\right) = -\frac{J'}{\sigma^2} + 2J\,\frac{\partial_y\sigma}{\sigma^3} = -\frac{J'}{\sigma^2} + \frac{2JQ}{\rho^2\sigma}.$$

Using $(5.27)_1$ and $(5.25)_b$ in (5.26) and comparing the result with $(5.27)_2$, we find

$$-\frac{J'}{\sigma^2}+\frac{2JQ}{\rho^2\sigma}=\frac{J'}{J}\left(-\frac{J}{\sigma^2}\right)-\frac{Q'\sigma}{\rho^3}-\left(-\frac{Q'\sigma}{\rho^3}+\frac{JQ}{\rho^2\sigma}-\frac{\omega}{\rho}\right)-\frac{3Q\sigma}{\rho^3}\left(-\frac{J}{\sigma^2}\right)$$
$$-\frac{1}{J}\left[\frac{\sigma^2}{\rho^2}\,\partial_x\omega+2\omega\sigma\left(-\frac{Q'\sigma}{Q\rho^2}+\frac{J}{\rho\sigma}-\frac{\omega}{Q}\right)+\frac{2\omega\sigma^2}{\rho}\left(-\frac{J}{\sigma^2}\right)\right],$$

which simplifies into

$$\frac{\omega}{\rho}-\frac{1}{J}\left(\frac{\sigma^2}{\rho^2}\,\partial_x\omega-\frac{2\omega\sigma^2Q'}{\rho^2Q}-\frac{2\omega^2\sigma}{Q}\right)=0,$$

that is,

$$(5.28)\qquad \frac{J\rho}{\sigma^2}-\frac{\partial_x\omega}{\omega}+\frac{2Q'}{Q}+\frac{2\omega\rho^2}{\sigma Q}=0.$$

This is the integrability condition we wished to establish. But because of $(5.18)_1$, $(5.24)_b$, we have

$$\frac{J\rho}{\sigma^2}=\frac{\partial_x\rho}{\rho}, \qquad \frac{\omega\rho^2}{\sigma Q}=\frac{\omega}{\sigma_y}=-\,\partial_x\log\,(\rho\sigma_y).$$

Therefore (5.28) becomes

$$\partial_x\log\left[\frac{\rho}{\omega}\left(\frac{Q}{\rho\sigma_y}\right)^2\right]=0,$$

which, by $(5.18)_{1a}$, can be written

$$(5.29)_a\qquad \partial_x\log\frac{\rho^3}{\omega\sigma^2}=0.$$

Following a procedure symmetric to the above one, we can prove that

$$(5.29)_b\qquad \partial_y\log\frac{\sigma^3}{\omega\rho^2}=0.$$

From (5.29) it follows at once that

$$(5.30)\qquad \partial_y\partial_x\log\frac{\rho}{\sigma}=0,$$

which show that ρ/σ is of the form $'\rho(y)/'\sigma(z)$. Thus the proof of Theorem 5.2 is completed.

Remark. The above proof for the case $n=m-n=2$ holds also for the more general case $n=m-n>1$, but not for other cases. Indeed, when $n>1$, $m-n>1$, equation $(5.17)_3$ may be replaced by (cf. (5.24))

$$\partial_y\rho_z + \rho_z\sigma_y = \frac{\omega}{n-1}, \qquad \partial_z\sigma_y + \rho_z\sigma_y = -\frac{\omega}{m-n-1},$$

and if we carry out on these equations a procedure similar to that which we did on (5.24) for the case $\omega\neq 0$, then the final result corresponding to (5.30) is

$$\partial_y\partial_z\left[\left(\frac{n-1}{m-n-1}+4\right)\log\rho - \left(\frac{m-n-1}{n-1}+4\right)\log\sigma\right] = 0.$$

This reduces to (5.30) when and only when $n=m-n$, which proves our assertion. Moreover, since our proof depends only on (5.17), (5.18), which are equivalent to (5.5), we have

If the fundamental tensor $^*g_{\alpha\beta}$ *of a* V_{2n} $(n>1)$ *is a properly conformally separable tensor of the type* (n, n) *and the Ricci tensors* $^*R_{\alpha\beta}$, $'R_{ij}$, $'R_{pq}$ *of* V_{2n} *and its subspaces* $x^r=$ *const. and* $x^k=$ *const. satisfy the relations:*

$$^*R_{ip} = 0, \qquad ^*R_{ij} - {'R_{ij}} \sim {^*g_{ij}}, \qquad ^*R_{pq} - {'R_{pq}} \sim {^*g_{pq}},$$

then $^*g_{\alpha\beta}$ *is conformal to a separable tensor of the type* (n, n).

6.1. **Main results.** Theorem 5.2 enables us to bring to a satisfactory conclusion our study of a properly conformally separable tensor which represents an E_m and each of whose component tensors either is of dimension 2 or represents E's.

Consider the properly conformally separable tensor

$$(6.1) \qquad ^*g_{\alpha\beta} = \tau^{-2}\,''g_{\alpha\beta} \equiv \tau^{-2}\begin{bmatrix}\bar g_{ij} & 0\\ 0 & \bar g_{pq}\end{bmatrix}, \qquad \begin{aligned}&\alpha, \beta, \gamma, \epsilon = 1, \cdots, m,\\ &i, j, k, l = 1, \cdots, n,\\ &p, q, r, s = n+1, \cdots, m,\end{aligned}$$

where

$$(6.2) \qquad \begin{gathered} n > 1, \qquad m-n > 1,\\ \tau = \tau(x^\alpha), \qquad \partial_i\tau \neq 0, \qquad \partial_p\tau \neq 0,\\ \bar g_{ij} = \bar g_{ij}(x^k), \qquad \bar g_{pq} = \bar g_{pq}(x^r).\end{gathered}$$

Let quantities referred to $''g_{\alpha\beta}$, $\bar g_{ij}$, or $\bar g_{pq}$ and covariant differentiations taken with respect to them be marked by the signs ($''$), ($^-$); (;), (/). Then we have (cf. (3.2), (3.3))

$$''\Gamma^k_{ij} = \bar\Gamma^k_{ij}, \qquad ''\Gamma^r_{pq} = \bar\Gamma^r_{pq}, \qquad ''\Gamma^\gamma_{\alpha\beta} = 0$$

if α, β, γ are not all in the same range; and consequently,

$$\tau_{:ip} = \partial_p \partial_i \tau, \qquad \tau_{:ij} = \tau_{/ij}, \qquad \tau_{:pq} = \tau_{/pq}, \tag{6.3}$$

$$''R_{ip} = 0, \qquad ''R_{ij} = \overline{R}_{ij}, \qquad ''R_{pq} = \overline{R}_{pq}. \tag{6.4}$$

Also from (6.1) we have [5, p. 90 (28.6)]

$$*R_{\alpha\beta} = ''R_{\alpha\beta} - (m-2)\frac{\tau_{:\alpha\beta}}{\tau} + ''g_{\alpha\beta}''g^{\gamma\epsilon}\left[-\frac{\tau_{:\gamma\epsilon}}{\tau} + (m-1)\frac{\tau_{:\gamma}\tau_{:\epsilon}}{\tau^2}\right]. \tag{6.5}$$

Now suppose that (6.1) represents an $E_m(c)$, that is, that

$$*R_{\alpha\beta} = -(m-1)c\,{}^*g_{\alpha\beta}, \qquad c = \text{const.} \tag{6.6}$$

In consequence of this and equations (6.1)–(6.4), equation (6.5) for $(\alpha, \beta) = (i, p)$ becomes $\partial_p \partial_i \tau = 0$, and therefore([16])

$$\tau = y(x^k) + z(x^r), \tag{6.7}$$

where $y \neq$ const., $z \neq$ const. because of (6.2). Hence (6.1) may be written

$$*g_{\alpha\beta} = (y+z)^{-2}\begin{bmatrix} \bar{g}_{ij} & 0 \\ 0 & \bar{g}_{pq} \end{bmatrix} \equiv (y+z)^{-2}\begin{bmatrix} y^2 g_{ij} & 0 \\ 0 & z^2 g_{pq} \end{bmatrix}. \tag{6.8}$$

If $n>2$ and the component tensor $*g_{ij} = [(y+z)/y]^{-2} g_{ij}$ represents E_n's (S_n's), then g_{ij} represents an E_n (S_n), because $*g_{ij}$ becomes g_{ij} for $z=0$. Consequently, we have by Theorem 2.1

$$\left(\frac{y+z}{y}\right)_{,ij} \sim g_{ij}, \quad \text{that is,} \quad \left(\frac{1}{y}\right)_{,ij} \sim g_{ij}, \tag{6.9}$$

where, as usual, a comma denotes covariant differentiation with respect to g_{ij}. Thus by Theorems 2.1 and 2.2, the tensor $y^2 g_{ij} = \bar{g}_{ij}$ is the fundamental tensor of an E_n (S_n).

We now suppose, besides (6.7), also that each of the component tensors of the tensor (6.8) either is of dimension 2 or represents E's. Then it follows from the above observation that

$$\overline{R}_{ij} = -(n-1)\bar{a}(x^k)\bar{g}_{ij}, \tag{6.10$_a$}$$

$$\overline{R}_{pq} = -(m-n-1)\bar{b}(x^r)\bar{g}_{pq}, \tag{6.10$_b$}$$

where $\bar{a}(x^k)$ is constant when $n>2$, and $\bar{b}(x^r)$ is constant when $m-n>2$. On account of (6.3), (6.4), (6.6), (6.7), (6.9) and (6.10), equation (6.5) for $(\alpha, \beta) = (i, j)$ and (p, q) become

$$\begin{aligned} -(m-1)c\tau^{-2}\bar{g}_{ij} = {} & -(n-1)\bar{a}(x^k)\bar{g}_{ij} - (m-2)\tau^{-1}y_{/ij} \\ & + \bar{g}_{ij}[-(\bar{g}^{hk}y_{/hk} + \bar{g}^{rs}z_{/rs})\tau^{-1} + (m-1)(\bar{g}^{hk}y_{/h}y_{/k} + \bar{g}^{rs}z_{/r}z_{/s})\tau^{-2}], \end{aligned} \tag{6.11$_a$}$$

([16]) We note that these y, z are not identical with the y, z which appeared in §§5.1–5.4.

$$(6.11)_b \quad \begin{aligned} -(m-1)c\tau^{-2}\bar{g}_{pq} &= -(m-n-1)\bar{b}(x^r)\bar{g}_{pq} - (m-2)\tau^{-1}z_{/pq} \\ &+ \bar{g}_{pq}[-(\bar{g}^{hk}y_{/hk} + \bar{g}^{rs}z_{/rs})\tau^{-1} + (m-1)(\bar{g}^{hk}y_{/h}y_{/k} + \bar{g}^{rs}z_{/r}z_{/s})\tau^{-2}]. \end{aligned}$$

From these it follows at once that

$$(6.12) \qquad y_{/ij} = -L\bar{g}_{ij}, \qquad z_{/pq} = -S\bar{g}_{pq},$$

which, because of (6.10), imply (cf. Theorems 2.3 and (2.9), (2.10)), respectively.

$$(6.13)_a \quad \bar{g}^{ij}y_{/i}y_{/j} = -2M(y), \quad L = L(y) = M'(y), \quad \bar{a}(x^k) = \bar{a}(y) = M''(y),$$

$$(6.13)_b \quad \bar{g}^{pq}z_{/p}z_{/q} = -2T(z), \quad S = S(z) = T'(z), \quad \bar{b}(x^r) = \bar{b}(z) = T''(z),$$

where the prime denotes differentiation. In consequence of these, equations (6.11) are equivalent to (6.12), (6.13) and

$$(6.14) \quad \begin{aligned} -(m-1)c\tau^{-2} + (n-1)M'' - (m-2)M'\tau^{-1} \\ = -(m-1)c\tau^{-2} + (m-n-1)T'' - (m-2)T'\tau^{-1} \\ = [nM' + (m-n)T']\tau^{-1} - 2(m-1)(M+T)\tau^{-2}. \end{aligned}$$

These last equations together with (6.7) can be solved for $M(y)$ and $T(z)$. Indeed, the first equation of (6.14) can be written

$$(6.15) \quad (m-2)(M'-T') - (y+z)[(n-1)M'' - (m-n-1)T''] = 0.$$

Differentiating this partially with respect to y twice, we find

$$(m-2n)M''' - (n-1)(y+z)M^{iv} = 0,$$

which, because M is a function of y alone, gives us $M^{iv}=0$. Hence

$$(6.16)_a \qquad M(y) = a_0 + a_1y + a_2y^2 + a_3y^3,$$

and by symmetry,

$$(6.16)_b \qquad T(z) = b_0 + b_1z + b_2z^2 + b_3z^3,$$

where the a's and b's are constants. When these values of M and T are substituted in (6.15), the latter reduces to

$$(6.17) \quad b_1 = a_1, \quad b_2 = -a_2, \quad \begin{cases} b_3 = a_3 = 0 & \text{if} \quad m \neq 2n, \\ b_3 = a_3 & \text{if} \quad m = 2n. \end{cases}$$

Now in consequence of (6.16) and (6.17), equations (6.14) become

$$(6.18) \qquad c = 2(a_0 + b_0),$$

$$(6.19) \qquad a_3 = 0 \quad \text{if} \quad m = 2n = 4 \text{ are not satisfied.}$$

Thus, the solution of equations (6.14) is given by (6.16)–(6.19).

By Theorem 5.2, the preceding results prove the necessity of the condition in the following

THEOREM 6.1. *In order that a properly conformally separable tensor* $^*g_{\alpha\beta}$ *of the type* $(n>1,\ m-n>1)$ *may represent an* E_m *and each of its component tensors either be of dimension* 2 *or represent E's, it is necessary and sufficient that* $^*g_{\alpha\beta}$ *be of the form* (6.8) *and equations* (6.10), (6.12), (6.13), (6.16)–(6.19) *be satisfied.*

The sufficiency of the condition in this theorem can be proved as follows. If $n>2$, $(6.10)_a$ shows that $\bar{g}_{ij}$ represents an E_n, and consequently by $(6.12)_a$, the tensor $g_{ij}=y^{-2}\bar{g}_{ij}$ also represents an E_n (Theorem 2.1). Thus equations (6.9) are satisfied, and accordingly the component tensor $^*g_{ij}=[(y+z)/y]^{-2}g_{ij}$ represents E_n's. Similarly, it follows from $(6.10)_b$ and $(6.12)_b$ that if $m-n>2$, the component tensor $^*g_{pq}$ represents E_{m-n}'s. Finally, from the way in which equations (6.10), (6.12), (6.13), (6.16)–(6.19) were derived, it is readily seen that if these equations are satisfied, equation (6.6) must also be satisfied. Hence $^*g_{\alpha\beta}$ represents an E_m and our theorem is completely proved.

Now we are ready to establish the following three main results.

THEOREM 6.2. *If a properly conformally separable tensor represents an* E_m *and one of its component tensors represents E's or* S_2*'s, then the other component tensor, if it is of dimension* 2, *represents* S_2*'s.*

THEOREM 6.3. *In order that a properly conformally separable tensor* $^*g_{\alpha\beta}$ *may represent an* E_m *and each of its component tensors E's or* S_2*'s it is necessary and sufficient that the following conditions be fulfilled:*

(1) $^*g_{\alpha\beta}$ *is of the form*

$$(y+z)^{-2}\begin{bmatrix}\bar{g}_{ij} & 0\\ 0 & \bar{g}_{pq}\end{bmatrix}, \qquad \begin{matrix} y=y(x^k), & z=z(x^r),\\ \bar{g}_{ij}=\bar{g}_{ij}(x^k), & \bar{g}_{pq}=\bar{g}_{pq}(x^r).\end{matrix}$$

(2) *The tensors* $\bar{g}_{ij}$, $\bar{g}_{pq}$ *each represent an E or* S_2 *with scalar curvatures* $\bar{a}$, $\bar{b}$ *connected by* $\bar{a}+\bar{b}=0$.

(3) *The equations*

$$y_{/ij}=-(\bar{a}y+f)\bar{g}_{ij}, \qquad z_{/pq}=-(\bar{b}z+f)\bar{g}_{pq}$$

are satisfied with a constant f.

If these conditions are fulfilled, the scalar curvature c of $^*g_{\alpha\beta}$ *is equal to the sum of the scalar curvatures of* $y^{-2}\bar{g}_{ij}$ *and* $z^{-2}\bar{g}_{pq}$ *(each of which, as is implied by* (2) *and* (3), *is the fundamental tensor of an E or* S_2).

THEOREM 6.4. *In order that a properly conformally separable tensor* $^*g_{\alpha\beta}$ *of the type* (2, 2), *whose component tensors do not represent* S_2*'s, may be the fundamental tensor of an* E_4, *it is necessary and sufficient that* (1) $^*g_{\alpha\beta}$ *be of the form*

$$A(\bar{a}+\bar{b})^{-2}\begin{bmatrix}\bar{g}_{ij} & 0\\ 0 & \bar{g}_{pq}\end{bmatrix}, \qquad A=\text{const.},$$

and (2) *the equations*

$$\bar{a}_{/ij} = -(1/2)(\bar{a}^2 + B)\bar{g}_{ij}, \qquad \bar{b}_{/pq} = -(1/2)(\bar{b}^2 + B)\bar{g}_{pq}$$

be satisfied with a constant B, where $\bar{a}$, $\bar{b}$ are the scalar curvatures of the fundamental tensors $\bar{g}_{ij}$, $\bar{g}_{pq}$.

The proof of these theorems will be based on Theorem 6.1. For Theorem 6.2, we suppose for the moment that $m-n=2$ and that if $n=2$, ${}^*g_{ij}$ represents S_2's. Theorem 6.2 will be proved if we can show that in this case $a_3=0$. Indeed, if $a_3=0$, then by (6.17), $b_3=0$, and consequently, by $(6.13)_b$ and $(6.16)_b$, $\bar{b}(x^r)=T''(z)=2b_2=\text{const}$. Therefore $\bar{g}_{pq}$ represents an S_2. Since $(6.12)_b$ can be written as $(y+z)_{/pq}=-S\bar{g}_{pq}$, the component tensor ${}^*g_{pq}=(y+z)^{-2}\bar{g}_{pq}$ represents S_2's. It remains therefore to prove that $a_3=0$.

If $n>2$, $a_3=0$ is given by (6.19) without further proof. If $n=2$, we have by hypothesis that ${}^*g_{ij}=(y+z)^{-2}\bar{g}_{ij}$ represents S_2's, which implies that the tensor $g_{ij}\equiv y^{-2}\bar{g}_{ij}$ also represents an S_2. Therefore, in consequence of $(6.12)_a$, $\bar{g}_{ij}=y^2g_{ij}$ is the fundamental tensor of an S_2 (Theorem 2.5). Hence we have from $(6.13)_a$ and $(6.16)_a$ that $\bar{a}(x^h)=M''(y)=2a_2+6a_3y=\text{const}$. From this it follows that $a_3=0$, as was to be proved. Theorem 6.1 has thus been completely established.

As a consequence of Theorem 6.2, for a properly conformally separable tensor which represents an E_m and each of whose component tensors either is of dimension 2 or represents E's, only two cases can happen: either (1) each of its component tensors represents E's or S_2's, or (2) $m=2n=2$ and neither of them represents S_2's. They are the two cases which we deal with in Theorems 6.3 and 6.4. For them, we have, respectively,

$$\begin{gathered} M(y) = a_0 + a_1y + a_2y^2, \qquad T(z) = b_0 + b_1z + b_2z^2, \\ a_1 = b_1, \qquad a_2 + b_2 = 0; \end{gathered} \tag{6.20}$$

$$\begin{gathered} M(y) = a_0 + a_1y + a_2y^2 + a_3y^3, \qquad T(z) = b_0 + b_1z + b_2z^2 + b_3z^3, \\ a_1 = b_1, \qquad a_2 + b_2 = 0, \qquad a_3 = b_3 \neq 0. \end{gathered} \tag{6.21}$$

Now for case (1), the scalar curvatures of the fundamental tensors $\bar{g}_{ij}$, $\bar{g}_{pq}$, $g_{ij}=y^{-2}\bar{g}_{ij}$, $g_{pq}=z^{-2}\bar{g}_{pq}$ are, respectively $\bar{a}=2a_2$, $\bar{b}=2b_2$, $2a_0$, $2b_0$, as follows from (6.13) and Theorem 2.4. Hence Theorem 6.3 is proved by (6.18) and the equations obtained by using (6.20) in (6.12) and (6.13).

Finally, to prove Theorem 6.4, we use (6.21), (6.13) in (6.12) and get

$$y_{/ij} = -(a_1 + 2a_2y + 3a_3y^2)\bar{g}_{ij}, \qquad z_{/pq} = -(a_1 - 2a_2z + 2a_3z^2)\bar{g}_{pq}. \tag{6.22}$$

Since $\bar{a}=2(a_2+3a_3y)$, $\bar{b}=2(-a_2+3a_3z)$ by (6.13), and $a_3\neq 0$, equations (6.7) and (6.22) can be expressed in terms of $\bar{a}$, $\bar{b}$. The result is readily found to be

$$\tau = A(\bar{a} + \bar{b}), \tag{6.23}$$

(6.24) $\quad \bar{a}_{/ij} = -(1/2)(\bar{a}^2 + B)\bar{g}_{ij}, \qquad \bar{b}_{/pq} = -(1/2)(\bar{b}^2 + B)\bar{g}_{pq},$

where $A = 1/(6a_3)$, $B = 4(a_2^2 - 3a_1a_3)$. This completes the proof of Theorems 6.2–6.4.

IV. Improperly conformally separable tensors

7.1. **A preliminary theorem.** The improperly conformally separable tensor with $\rho = 1$ is

$$^*g_{\alpha\beta} = \begin{bmatrix} g_{ij} & 0 \\ 0 & \sigma^{-2}g_{pq} \end{bmatrix}, \qquad \begin{aligned} &\alpha, \beta, \gamma, \epsilon = 1, \cdots, m, \\ &i, j, k, l = 1, \cdots, n, \\ &p, q, r, s = n + 1, \cdots, m, \end{aligned} \tag{7.1}$$

where

$$\sigma = \sigma(x^a), \qquad g_{ij} = g_{ij}(x^k), \qquad g_{pq} = g_{pq}(x^r).$$

We suppose throughout this section that $\partial_i\sigma \neq 0$, that is, that $^*g_{\alpha\beta}$ is not separable in the ordinary sense. For the tensor (7.1) we have (cf. (3.6))

$$\begin{aligned} ^*R_{ip} &= -(m - n - 1)\partial_p\sigma_i, \\ ^*R_{ij} &= R_{ij} + (m - n)\sigma\left(\frac{1}{\sigma}\right)_{,ij}, \\ ^*R_{pq} &= {}'R_{pq} + {}^*g_{pq}g^{ij}\left[\sigma\left(\frac{1}{\sigma}\right)_{,ij} + (m - n - 1)\sigma^2\left(\frac{1}{\sigma}\right)_{,i}\left(\frac{1}{\sigma}\right)_{,j}\right], \end{aligned} \tag{7.2}$$

where the comma denotes covariant differentiation with respect to g_{ij}. Let (7.1) represent an $E_m(c)$, so that

$$^*R_{\alpha\beta} = -(m - 1)c\,^*g_{\alpha\beta}, \qquad c = \text{const.} \tag{7.3}$$

On account of this, $(7.2)_1$ becomes

$$(m - n - 1)\partial_p\sigma_i = 0.$$

Hence

THEOREM 7.1. *If the improperly conformally separable tensor* (7.1) *represents an E_m, then either $m = n+1$ or σ is of the form $z(x^r)/y(x^k)$.*

We discuss these two cases separately.

7.2. **The case $m = n+1$.** Since $\pm g_{mm}$ may be absorbed in σ^{-2}, (7.1) may be written

$$^*g_{\alpha\beta} = \begin{bmatrix} g_{ij} & 0 \\ 0 & e\sigma^{-2} \end{bmatrix}, \qquad e = \pm 1. \tag{7.4}$$

For this case $'R_{mm} = 0$, and, in consequence of (7.2), equation (7.3) is equivalent to

$$
(7.5)\qquad
\begin{aligned}
R_{ij} + \sigma\left(\frac{1}{\sigma}\right)_{,ij} &= -\,ncg_{ij},\\
g^{ij}\sigma\left(\frac{1}{\sigma}\right)_{,ij} &= -\,nc.
\end{aligned}
$$

If $n=2$, we have shown (Theorem 4.1) that g_{ij} represents an S_2. We now consider the more general case when $n>1$ and the component tensor g_{ij} is such that

$$
(7.6)\qquad R_{ij} = -\,(n-1)ag_{ij}, \qquad a = \text{const.},
$$

that is, that g_{ij} represents an E_n or an S_2. Because of (7.6), equation $(7.5)_1$ becomes

$$
(7.7)\qquad \left(\frac{1}{\sigma}\right)_{,ij} = -\,[nc-(n-1)a]\frac{1}{\sigma}g_{ij}.
$$

Transvecting this by g^{ij} and comparing the result with $(7.5)_2$, we find that

$$
(7.8)\qquad a = c.
$$

On account of this, equation (7.7) becomes

$$
(7.9)\qquad \left(\frac{1}{\sigma}\right)_{,ij} = -\frac{a}{\sigma}g_{ij}.
$$

Since $(7.5)_2$ is evidently a consequence of (7.8) and (7.9), the latter equations, because of (7.6), are equivalent to (7.5), and hence to (7.3). Thus we have proved the following

THEOREM 7.2 (i). *Let g_{ij} be the fundamental tensor of an $E_n(a)$ or $S_2(a)$. Then in order that the improperly conformally separable tensor (7.4) of the type ($n>1$, $m-n=1$) may represent an E_m, it is necessary and sufficient that g_{ij} and σ satisfy equation (7.9).*

Since by supposition g_{ij} represents an $E_n(a)$ or $S_2(a)$, it follows from (7.9) that the tensor $\sigma^2 g_{ij}$ represents E_n's or S_2's (Theorem 2.1). If, in particular, $a=0$, then (7.9) reduces to $(1/\sigma)_{,ij}=0$, and therefore the $E_n(0)$ or $S_2(0)$ with fundamental tensor g_{ij} has a parallel vector field. Hence

THEOREM 7.2 (ii). *If the improperly conformally separable tensor (7.4) and its first component tensor g_{ij} represent an $E_m(c)$ and an $E_n(a)$ or $S_2(a)$, respectively, then the tensor $\sigma^2 g_{ij}$ represents E_n's or S_2's and $c=a$. If $c=a=0$, then the $E_n(0)$ or $S_2(0)$ with fundamental tensor g_{ij} possesses a parallel vector field.*

Conversely, let g_{ij} be the fundamental tensor of an $E_n(a)$ which is conformal[17] to another E_n, then the equation

[17] An E_n with fundamental tensor g_{ij} is said to be conformal to another E_n if a nonconstant scalar y exists such that $y^{-2}g_{ij}$ is the fundamental tensor of an E_n.

$$y_{,ij} = -(ay+f)g_{ij}, \qquad f = \text{const.}$$

has a solution for y (cf. Theorems 2.1, 2.3 and (2.11)). Thus, if $a \neq 0$, the function $1/\sigma = y + f/a$ evidently satisfies (7.9). Hence we have the following converse to Theorem 7.2 (ii):

THEOREM 7.2 (iii). *Given an E_n whose scalar curvature is zero and which has a parallel vector field, or one whose scalar curvature is not zero and which is conformal to another E_n, then the given E_n can be isometrically imbedded in an E_{n+1} as a member of ∞^1 isometric, non-parallel, and totally geodesic hypersurfaces.*

The present case has already been considered with a different method by Fialkow [7, 7′], and the results stated in Theorems 7.2 (i), (ii) and (iii) are due to him. However(18), he overlooked the exceptional case $E_n(0)$, for which the property of its being conformal to another E_n is not a sufficient condition for it to be imbeddable in an E_{n+1} in the manner stated in Theorem 7.2 (iii).

7.3. **The case $m-n>1$.** By Theorem 7.1, in this case σ must be of the form $z(x^r)/y(x^k)$. Since $z(x^r)$ may be absorbed in g_{pq}, there is no loss of generality in assuming that $\sigma = 1/y(x^k)$. Thus the conformally separable tensor under consideration takes the form

$$(7.10) \qquad {}^*g_{\alpha\beta} = \begin{bmatrix} g_{ij} & 0 \\ 0 & y^2 g_{pq} \end{bmatrix},$$

where $y \neq$ const. On account of (7.3) and the fact that the fundamental tensors g_{pq} and ${}^*g_{pq} = y^2 g_{pq}$ have identical Ricci tensors, equation $(7.2)_3$ for the tensor (7.10) can be written

$$(7.11) \quad R_{pq} = {}'R_{pq} = -\left\{(m-1)c + g^{ij}\left[\frac{y_{,ij}}{y} + (m-n-1)\frac{y_{,i}y_{,j}}{y^2}\right]\right\} y^2 g_{pq}.$$

(18) Equations (3.11) of Fialkow [7] represent a necessary and sufficient condition for an E_n to be imbeddable in an E_{n+1} as a member of ∞^1 isometric, non-parallel, totally geodesic hypersurfaces. In p. 427 (line 18) of the same paper, we find the sentence "According to Brinkmann (3.11) is the necessary and sufficient condition that E_n be conformal to another Einstein space by means of a transformation $d\bar{s} = \sigma ds$ with $\Delta_1\sigma \neq 0$, where $\Delta_1\sigma = f^{ij}\sigma_{,i}\sigma_{,j}$." This sentence is not entirely correct; in fact, the condition is sufficient but *not necessary*. To explain, we use Fialkow's notation. Brinkmann's original necessary and sufficient condition referred to above is [3, p. 125, Theorem II] that a coordinate system exists in which the fundamental tensor of E_n is of the form

$$(\Delta) \qquad \begin{gathered} f_{nn} = (cx^{n^2} + 2Ax^n + d)^{-1}, \\ f_{\lambda\mu} = (cx^{n^2} + 2Ax^n + d)F_{\lambda\mu}(x^\nu), \qquad f_{n\lambda} = 0, \\ A \text{ and } d \text{ constant,} \end{gathered}$$

and the form $F_{\lambda\mu}(x^\nu)\,dx^\lambda dx^\mu$ is the fundamental form of an E_{n-1}. Equations (3.11) of Fialkow [7] differ from (Δ) by the absence of the constant A, and, (Δ) can be reduced, by putting $\bar{x}^n = x^n + B$, $B =$ const., to (3.11) of Fialkow [7] when $c \neq 0$ but not when $c = 0$. This justifies our statement.

Hence

THEOREM 7.3. *If the improperly conformally separable tensor* (7.10) *of the type* $(n, m-n>2)$ *represents an* E_m, *then its second component tensor* $^*g_{pq}=y^2g_{pq}$ *represents* E_{m-n}*'s.*

We shall now consider separately the following three subcases: (1) $n=1$; (2) $n>1$ and g_{ij} represents an E_n or S_2; (3) $n=2$.

Subcase 1. $n=1$, $m-n>1$.

The conformally separable tensor in question is

$$(7.12) \qquad {}^*g_{\alpha\beta} = \begin{bmatrix} g_{11} & 0 \\ 0 & y^2g_{pq} \end{bmatrix}, \qquad y = y(x^1) \neq \text{const.}$$

With n, x, g_{11}, y^2g_{pq} in place of $m-n$, x^m, $e\sigma^{-2}$, $\rho^{-2}g_{ij}$, respectively, the case considered in §4 reduces to the present subcase. Hence from Theorems 4.1 and 4.2, we have

THEOREM 7.4. *In order that the improperly conformally separable tensor* (7.12) *with* $m>2$ *may represent an* $E_m(c)$, *it is necessary and sufficient that* g_{pq} *represent an* $E_{m-1}(b)$ *or* $S_2(b)$ *and*

$$(7.13) \qquad g_{11} = \frac{(y')^2}{b - cy^2}.$$

This result can also be proved directly from (7.2) and (7.3).

Subcase 2. $n>1$, $m-n>1$, *and* g_{ij} *represents an* E_n *or* S_2.

By hypothesis, we have, besides (7.2), (7.3), (7.11), also

$$(7.14) \qquad R_{ij} = -(n-1)ag_{ij}, \qquad a = \text{const.}$$

In consequence of this and (7.3), equation $(7.2)_2$ becomes

$$(7.15) \qquad y_{,ij} = -\frac{(m-1)c-(n-1)a}{m-n}\, yg_{ij}.$$

This equation is of the form (2.1). And since (7.14) is satisfied, it follows from $(2.11)_1$ that (cf. Theorem 2.3)

$$\frac{(m-1)c-(n-1)a}{m-n}\, y = ay + f, \qquad f = \text{const.},$$

which give $f=0$ and $a=c$. Therefore (7.15) is equivalent to

$$(7.16) \qquad y_{,ij} = -ayg_{ij},$$

$$(7.17) \qquad c = a.$$

As a consequence of (7.14) and (7.16), $\bar{g}_{ij}=y^{-2}g_{ij}$ is the fundamental tensor of an $E_n(\bar{a})$ or $S_2(\bar{a})$, where $\bar{a}$ is determined from (cf. Theorem 2.4)

$$g^{ij}y_{,i}y_{,j} = -(ay^2 + \bar{a}). \tag{7.18}$$

When (7.16)–(7.18) are used in (7.11), the latter becomes

$$R_{pq} = -(m - n - 1)(-\bar{a})g_{pq}, \tag{7.19}$$

which shows that g_{pq} represents an $E_{m-n}(-\bar{a})$ or $S_2(-\bar{a})$. Hence

THEOREM 7.5 (i). *If the improperly conformally separable tensor* (7.10) *of the type* ($n>1$, $m-n>1$) *and its component tensor g_{ij} represent, respectively, an $E_m(c)$ and an $E_n(a)$ or $S_2(a)$, then* (1) *g_{pq} represents an $E_{m-n}(b)$ or $S_2(b)$,* (2) *$y^{-2}g_{ij}$ represents an $E_n(\bar{a})$ or $S_2(\bar{a})$, and* (3) *$c=a$, $b=-\bar{a}$. If $c=a=0$, the $E_n(0)$ or $S_2(0)$ with fundamental tensor g_{ij} possesses a parallel vector field.*

We observe that when (7.14) is supposed, equations (7.2), (7.3) are equivalent to (7.16)–(7.19). Then, by a consideration similar to that leading to Theorem 7.2 (iii), we can prove the following

THEOREM 7.5 (ii). *Given an E_n whose scalar curvature is zero and which possesses a parallel vector field or one whose scalar curvature is not zero and which is conformal to another E_n, then the given E_n can be isometrically imbedded in an E_m of any dimension m greater than $n+1$ as a member of ∞^{m-n} isometric and totally geodesic subspaces E_n's which are orthogonal to ∞^n totally umbilical subspaces E_{m-n}'s or S_2's.*

Subcase 3. $n=2$, $m-n>1$.

For this case we have

$$R_{ij} = -a(x^k)g_{ij}. \tag{7.20}$$

In consequence of this and (7.3), equation $(7.2)_2$ becomes

$$y_{,ij} = -\frac{(m-1)c - a(x^k)}{m-2}\,yg_{ij}. \tag{7.21}$$

This is of the form (2.1), and consequently, it follows from (2.9) that (cf. Theorem 2.3)

$$a(x^k) = a(y) = \frac{d}{dy}\left[\frac{(m-1)c - a(y)}{m-2}\,y\right],$$

that is,

$$a'(y)y + (m-1)a(y) = (m-1)c.$$

Multiplying this by y^{m-2} and then integrating, we have

$$a(y) = c - (m-2)fy^{1-m}, \qquad f = \text{const.} \tag{7.22}$$

Here it is evident that $a(y)$ is constant or not according as f is or is not zero.

With this value of $a(y)$, equation (7.21) becomes

$$y_{,ij} = -(cy + fy^{2-m})g_{ij}. \tag{7.23}$$

From this and (2.10), which holds by Theorem 2.3, it follows that

$$g^{ij}y_{,i}y_{,j} = -\left(cy^2 + \frac{2f}{3-m}y^{3-m} - b\right), \qquad b = \text{const.} \tag{7.24}$$

When (7.23) and (7.24) are substituted in (7.11) for $n=2$, the result is

$$R_{pq} = -(m-3)bg_{pq}, \tag{7.25}$$

which is the condition for g_{pq} to represent an $E_{m-2}(b)$ or $S_2(b)$. Hence

THEOREM 7.6. *In order that the improperly conformally separable tensor* (7.10) *of the type* ($n=2$, $m-n>1$) *may represent an* $E_m(c)$, *it is necessary and sufficient that* (1) g_{ij} *and* y *be such that equation* (7.23) *is satisfied with a constant* f, *and* (2) g_{pq} *represent an* E_{m-2} *or* S_2 *of scalar curvature* b *given by* (7.24).

As a verification we observe that the result for subcase 2 with $n=2$ is identical with the result for subcase 3 with $a=$const. (that is, $f=0$).

Finally, it follows from the last but one paragraph of §2 that a 2-dimensional fundamental tensor g_{ij} actually exists whose scalar curvature is not constant and for which equation (7.23) admits a solution for y. Thus the existence of an improperly conformally separable tensor which represents an E_m and whose component tensors g_{ij} and $^*g_{pq}$ are such that, the first represents a V_2 which is not an S_2, while the second represents E_{m-2}'s or S_2's. This fact is in contrast with Theorem 6.2 of the preceding section and Theorem 8.1 of the following section.

8. **Separable tensors.** In this section we reproduce some results of Fialkow concerning a separable tensor, thus completing our discussion of the conformally separable tensor which represents an E_m and each of whose component tensors either is of dimension less than 3 or represents E's.

For the separable tensor

$$^*g_{\alpha\beta} = \begin{bmatrix} g_{ij} & 0 \\ 0 & g_{pq} \end{bmatrix}, \tag{8.1}$$

equations (3.6) reduce to

$$^*R_{ip} = 0, \qquad ^*R_{ij} = R_{ij}, \qquad ^*R_{pq} = R_{pq}. \tag{8.2}$$

Suppose that (8.1) represents an $E_m(c)$, that is, that

$$^*R_{\alpha\beta} = -(m-1)c^*g_{\alpha\beta}, \qquad c = \text{const.} \tag{8.3}$$

Then according as $n>1$, $m-n>1$, or $n>1$, $m-n=1$, (8.2) become

$$0 = 0, \qquad R_{ij} = -(m-1)cg_{ij}, \qquad R_{pq} = -(m-1)cg_{pq}; \tag{8.4}$$

or

$$(8.5) \qquad 0 = 0, \qquad R_{ij} = -(m-1)cg_{ij}, \qquad 0 = -(m-1)cg_{mm}.$$

From these we have at once the following results due to Fialkow [7]:

THEOREM 8.1. *A separable tensor of the type $(n>1,\ m-n>1)$ represents an $E_m(c)$, if and only if its component tensors g_{ij} and g_{pq} represent, respectively, an $E_n(a)$ or $S_2(a)$ and an $E_{m-n}(b)$ or $S_2(b)$, and $(m-1)c=(n-1)a=(m-n-1)b$.*

THEOREM 8.2. *A separable tensor of the type $(n>1,\ m-n=1)$ represents an $E_m(c)$, if and only if its component tensor g_{ij} represents an $E_n(0)$ or $S_2(0)$. Then $c=0$.*

V. PARTICULAR CASES

9. S_m with conformally separable fundamental tensor. It is well known that the fundamental form of an $S_m(c)$ can always be reduced to the Riemann form [5, p. 85]

$$(9.1) \qquad \frac{e_1(dx^1)^2 + \cdots + e_m(dx^m)^2}{\{1 + (c/4)[e_1(x^1)^2 + \cdots + e_m(x^m)^2]\}^2},$$

where each e is ± 1. The form (9.1) is evidently properly conformally separable(19) of the type $(n,\ m-n)$, where n may be any integer from 1 to $m-1$. In what follows we give a few theorems concerning a conformally separable tensor which represents an S_m. Throughout this section all the symbols and indices have the same usual meaning.

From (3.4) it follows that the condition

$$(9.2) \qquad {}^*R^{\epsilon}_{\alpha\beta\gamma} = c(\delta^{\epsilon}_{\beta}\,{}^*g_{\alpha\gamma} - \delta^{\epsilon}_{\gamma}\,{}^*g_{\alpha\beta})$$

for the conformally separable tensor (1.1) to represent an S_m becomes

$$(9.3) \qquad \begin{gathered} {}'R^{l}_{ijk} = (c + \rho^{p}\rho_p)(\delta^{l}_{j}\,{}^*g_{ik} - \delta^{l}_{k}\,{}^*g_{ij}), \\ {}'R^{s}_{pqr} = (c + \sigma^{i}\sigma_i)(\delta^{s}_{q}\,{}^*g_{pr} - \delta^{s}_{r}\,{}^*g_{pq}), \\ (n-1)(\partial_i\rho_p + \rho_p\sigma_i) = 0, \qquad (m-n-1)(\partial_p\sigma_i + \rho_p\sigma_i) = 0, \\ \rho\left(\frac{1}{\rho}\right)_{;pq} = -\left[c + \frac{1}{n}\,{}^*g^{ij}\sigma\left(\frac{1}{\sigma}\right)_{;ij}\right]{}^*g_{pq}, \\ \sigma\left(\frac{1}{\sigma}\right)_{;ij} = -\left[c + \frac{1}{m-n}\,{}^*g^{pq}\rho\left(\frac{1}{\rho}\right)_{;pq}\right]{}^*g_{ij}. \end{gathered}$$

For $\rho=\sigma=1$, we have immediately [9, p. 896]:

THEOREM 9.1. *A separable tensor cannot represent an S_m of nonzero scalar curvature.*

(19) The fundamental form ${}^*g_{\alpha\beta}dx^{\alpha}dx^{\beta}$ is said to be conformally separable, if the tensor $g_{\alpha\beta}$ is conformally separable, and vice versa.

From $(9.3)_1$ and $(9.3)_2$ we have the following theorem [14]:

THEOREM 9.2. *If a conformally separable tensor represents an S_m, then each of its component tensors, if it is of dimension greater than 2, represents S's.*

Since an S is necessarily an E, it follows from this theorem that all properties of the conformally separable tensor which represents an E_m and each of whose component tensors either is of dimension less than 3 or represents E's must also be shared by the conformally separable tensor which represents an S_m. This observation together with what we have obtained in §§4–7 enables us to prove the following extension of Theorem 9.2:

THEOREM 9.3. *If a conformally separable tensor ${}^*g_{\alpha\beta}$ which is not an ordinarily separable tensor represents an S_m, then each of its component tensors either is of dimension 1 or represents S's. Conversely, if ${}^*g_{\alpha\beta}$ represents an E_m and each of its component tensors either is of dimension 1 or represents S's, then the E_m is an S_m.*

Proof. The proof of this theorem has to be carried out separately for several cases. Consider first the case of the properly conformally separable tensor, which we subdivide into the following three types: (1) $n=2$, $m-n=1$, (2) $n>2$, $m-n=1$, (3) $n>1$, $m-n>1$.

For the type ($n=2$, $m-n=1$), Theorem 9.3 reduces to Theorem 4.1 for $m=2$, because an E_3 is an S_3.

For the type ($n>2$, $m-n=1$), equations (9.3) become

$$\begin{aligned} {}'R^{l}_{ijk} &= (c + \overset{m}{\rho}\rho_m)(\delta^{l}_{j}\,{}^*g_{ik} - \delta^{l}_{k}\,{}^*g_{ij}), \\ \partial_i\rho_m + \rho_m\sigma_i &= 0, \\ \rho\left(\frac{1}{\rho}\right)_{;mm} &= -\left[c + \frac{1}{n}\,{}^*g^{ij}\sigma\left(\frac{1}{\sigma}\right)_{;ij}\right]{}^*g_{mm}, \\ \sigma\left(\frac{1}{\sigma}\right)_{;ij} &= -\left[c + {}^*g^{mm}\rho\left(\frac{1}{\rho}\right)_{;mm}\right]{}^*g_{ij}. \end{aligned} \tag{9.4}$$

The first part of the theorem follows at once from $(9.4)_1$. To prove the converse part, we suppose that ${}^*g_{\alpha\beta}$ and its component tensor ${}^*g_{ij}=\rho^{-2}g_{ij}$ represent, respectively, an E_m and S_n's, and show that (9.4) are satisfied. From this supposition it follows that all the equations appearing in §4 (with the exception of (4.6′)) and the equation

$$ {}'R^{l}_{ijk} = {}^*a(\overset{m}{x})(\delta^{l}_{j}\,{}^*g_{ik} - \delta^{l}_{k}\,{}^*g_{ij}) \tag{9.5}$$

are true. Now equations $(9.4)_2$ and $(9.4)_3$ are satisfied because of $(4.2)_1$, $(4.2)_2$ and (4.3). And on account of (9.5), equation $(9.4)_1$ becomes

$$c = {}^*a(x^m) - {}^*g^{mm}(\rho_m)^2,$$

which, by Theorem 4.1, is identical with (4.6). Finally, by virtue of (4.15), equation $(9.4)_4$ becomes (4.14′). Thus equations (9.4) are satisfied, as was to be proved.

We can now come to the properly conformally separable tensor of the type $(n>1,\ m-n>1)$. As a consequence of Theorems 9.2 and 5.2, a properly conformally separable tensor of this type which represents an S_m must be of the form

$$(9.6)\qquad {}^*g_{\alpha\beta} = \tau^{-2}\begin{bmatrix} \bar{g}_{ij} & 0 \\ 0 & \bar{g}_{pq} \end{bmatrix}.$$

The Riemann tensors ${}^*R^{\epsilon}_{\alpha\beta\gamma}$ and ${}''R^{\epsilon}_{\alpha\beta\gamma}$ of the tensors ${}^*g_{\alpha\beta}$ and

$$(9.7)\qquad {}''g_{\alpha\beta} = \begin{bmatrix} \bar{g}_{ij} & 0 \\ 0 & \bar{g}_{pq} \end{bmatrix}$$

are connected by (cf. (2.5))

$$(9.8)\qquad \begin{aligned} {}^*R^{\epsilon}_{\alpha\beta\gamma} = {}''R^{\epsilon}_{\alpha\beta\gamma} &+ (\delta^{\epsilon}_{\beta}\tau_{:\alpha\gamma} - \delta^{\epsilon}_{\gamma}\tau_{:\alpha\beta})\tau^{-1} + {}''g^{\kappa\epsilon}(\tau_{:\kappa\beta}{}''g_{\alpha\gamma} - \tau_{:\kappa\gamma}{}''g_{\alpha\beta})\tau^{-1} \\ &- {}''g^{\kappa\lambda}\tau_{:\kappa}\tau_{:\lambda}(\delta^{\epsilon}_{\beta}{}''g_{\alpha\gamma} - \delta^{\epsilon}_{\gamma}{}''g_{\alpha\beta})\tau^{-2}, \end{aligned}$$

where the colon denotes covariant differentiation with respect to ${}''g_{\alpha\beta}$; the indices $\alpha, \beta, \gamma, \epsilon, \kappa, \lambda$ have the range $1, \cdots, m$; and the components of ${}''R^{\epsilon}_{\alpha\beta\gamma}$ are (cf. (9.7), (3.4))

$$(9.9)\qquad \begin{aligned} &{}''R^{l}_{ijk} = \bar{R}^{l}_{ijk}, \qquad {}''R^{s}_{pqr} = \bar{R}^{s}_{pqr}, \\ &{}''R^{\epsilon}_{\alpha\beta\gamma} = 0 \qquad \text{if } \alpha, \beta, \gamma, \epsilon \text{ are not all in the same range.} \end{aligned}$$

From Theorem 9.2 and the remark below it and also the equations in §6, it follows that if the tensor ${}^*g_{\alpha\beta}$ defined by (9.6) represents an $S_m(c)$, then we have

$$(9.10)\qquad \begin{aligned} &\tau = y(x^h) + z(x^r), \\ &M = a_0 + a_1 y + a_2 y^2 + a_3 y^3, \qquad T = b_0 + a_1 z - a_2 z^2 + a_3 z^3, \\ &y_{/ij} = -M'\bar{g}_{ij}, \qquad z_{/pq} = -T'\bar{g}_{pq}, \\ &\bar{R}^{l}_{ijk} = M''(\delta^{l}_{j}\bar{g}_{ik} - \delta^{l}_{k}\bar{g}_{ij}), \qquad \bar{R}^{s}_{pqr} = T''(\delta^{s}_{q}\bar{g}_{pr} - \delta^{s}_{r}\bar{g}_{pq}), \end{aligned}$$

where $a_3=0$ unless $m=2n=4$. It can now be readily verified that in consequence of (9.8)–(9.10), equation (9.2) becomes $a_3=0$. On account of this, it follows easily from (9.10) that if any of the component tensors $\tau^{-2}\bar{g}_{ij}$ and $\tau^{-2}\bar{g}_{pq}$ is of dimension 2, then it represents S_2's. The first part of Theorem 9.3 for the present case is thus proved. We now suppose that ${}^*g_{\alpha\beta}$ represents an E and its component tensors S's. Then (9.10) are satisfied with $a_3=0$, and

consequently, equation (9.2) is satisfied, as follows from the sentence below (9.10). Hence the E_m in question is an S_m, and the proof of our theorem for the case of a properly conformally separable tensor is completed.

Finally, for an improperly conformally separable tensor with $\rho=1, \partial_i\sigma\neq 0$, equations (9.3) become

$$
\begin{aligned}
&R^{l}_{ijk} = c(\delta^{l}_{j}g_{ik} - \delta^{l}_{k}g_{ij}),\\
&{}'R^{s}_{pqr} = (c + \sigma^{i}\sigma_{i})(\delta^{s}_{q}{}^{*}g_{pr} - \delta^{s}_{r}{}^{*}g_{pq}),\\
&(m - n - 1)\partial_{p}\sigma_{i} = 0,\\
&\sigma\left(\frac{1}{\sigma}\right)_{,ij} = - c g_{ij}.
\end{aligned}
\tag{9.11}
$$

If we compare these equations with those appearing in §7, we shall see easily that Theorem 9.3 is true. Theorem 9.3 has thus been completely proved.

10. E_4 **with conformally separable fundamental tensor of the type** (2, 2). Let us review the cases we have considered for a conformally separable tensor which represents an E_4.

The type (3, 1). We have considered only the particular case when the first component tensor g_{ij} represents E_3's, that is, S_3's. By Theorem 9.3, an E_4 with such a conformally separable fundamental tensor is an S_4.

The type (2, 2). If ${}^{*}g_{\alpha\beta}$ is a properly conformally separable tensor, then either its component tensors both represent S_2's or neither of them does (cf. Theorem 6.2); for the respective cases, the E_4 is an S_4 or not an S_4 (cf. Theorem 9.3). If ${}^{*}g_{\alpha\beta}$ is an improperly conformally separable tensor with $\rho=1$, $\partial_i\sigma\neq 0$, the component tensor ${}^{*}g_{pq}$ necessarily represents S_2's, while the other, g_{ij}, may or may not (cf. Theorem 7.6); for the respective cases, the E_4 is an S_4 or not an S_4. Finally, if ${}^{*}g_{\alpha\beta}$ is a separable tensor, its two component tensors each represent an S_2 of equal nonzero scalar curvature (cf. Theorems 8.1 and 9.1).

Hence we have three and only three cases in which the E_4 is not an S_4; they are the cases of Theorem 6.4, Theorem 7.6 for $m=4$ and $f\neq 0$, and Theorem 8.1 for $m=2n=4$ and $a\neq 0$. From these theorems and (2.14), (9.1) we have at once the following

THEOREM 10.1. *A conformally separable fundamental form of the type* (2, 2) *represents an E_4 which is not an S_4, if and only if it can be reduced to one of the following forms* (*in which each e is* ± 1; *A, B, C, D are constants, and* $A\neq 0$):

$$
\begin{aligned}
&A(x^1 + x^3)^{-2}\left\{\frac{(dx^1)^2}{(1/3)(x^1)^3 + Bx^1 + C} + e_2[(1/3)(x^1)^3 + Bx^1 + C](dx^2)^2\right.\\
&\left.\qquad + \frac{(dx^3)^2}{(1/3)(x^3)^3 + Bx^3 + D} + e_4[(1/3)(x^3)^3 + Bx^3 + D](dx^4)^2\right\},
\end{aligned}
\tag{1}
$$

$$(2)\qquad \frac{(dx^1)^2}{C(x^1)^2+A(x^1)^{-1}+B}+e_2[C(x^1)^2+A(x^1)^{-1}+B](dx^2)^2 + \frac{(x^1)^2[e_3(dx^3)^2+e_4(dx^4)^2]}{[1+(B/4)(e_3(x^3)^2+e_4(x^4)^2)]^2},$$

$$(3)\qquad \frac{e_1(dx^1)^2+e_2(dx^2)^2}{[1+(A/4)(e_1(x^1)^2+e_2(x^2)^2)]^2}+\frac{e_3(dx^3)^2+e_4(dx^4)^2}{[1+(A/4)(e_3(x^3)^2+e_4(x^4)^2)]^2}.$$

Form (3) has been obtained by Kasner [11], and a form which is essentially the same as (2) for $B=1$ by Kottler [12, p. 443]. Form (1) however seems to be introduced here for the first time.

References

1. E. Bompiani, *Spazi Riemanniani luoghi di totalmente geodetiche*, Rend. Circ. Mat. Palermo vol. 48 (1924) pp. 121–134.

2. H. W. Brinckmann, *Riemann spaces conformal to Einstein spaces*, Math. Ann. vol. 91 (1924) pp. 269–278.

3. ———, *Einstein spaces which are mapped conformally on each other*, ibid. vol. 94 (1925) pp. 119–145.

4. A. Delgleize, *Sur les equations de Weingarten et les espaces pseudosphériques*, Bulletin de la Societé royale des Sciences de Liège vol. 4 (1935) pp. 158–161.

5. L. P. Eisenhart, *Riemannian geometry*, 1926, Princeton.

6. A. Fialkow, *Einstein spaces in a space of constant curvature*, Proc. Nat. Acad. Sci. U.S.A. vol. 24 (1938) pp. 30–34.

7. ———, *Totally geodesic Einstein spaces*, Bull. Amer. Math. Soc. vol. 45 (1939) pp. 423–428.

7′. ———, *Correction to "Totally geodesic Einstein spaces,"* ibid. vol. 48 (1942) pp. 167–168.

8. ———, *Conformal geodesics*, Trans. Amer. Math. Soc. vol. 45 (1939) pp. 443–473.

9. F. A. Ficken, *The Riemannian and affine differential geometry of product-spaces*, Ann. of Math. (2) vol. 40 (1939) pp. 892–913.

10. E. Kasner, *An algebraic solution of the Einstein equations*, Trans. Amer. Math. Soc. vol. 27 (1925) pp. 101–105.

11. ———, *Separable quadratic differential forms and Einstein solutions*, Proc. Nat. Acad. Sci. U.S.A. vol. 11 (1925) pp. 95–96.

12. F. Kottler, *Über die physikalischen Grundlagen der Einsteinsch Gravitationstheorie*, Annalen der Physik (4) vol. 56 (1918) pp. 401–462.

13. J. A. Schouten, and D. J. Struik, *On some properties of general manifolds relating to Einstein's theory of gravitation*, Amer. J. Math. vol. 43 (1921) pp. 217–221.

14. K. Yano, *Conformally separable quadratic differential forms*, Proc. Imp. Acad. Tokyo vol. 16 (1940) pp. 83–86.

15. ———, *Concircular geometry*, I. *Concircular transformations*, loc. cit. pp. 195–200.

16. ———, *Concircular geometry*, II. *Integrability conditions of* $\rho_{\mu\nu}=\phi g_{\mu\nu}$, loc. cit. pp. 354–360.

17. Yung-Chow Wong, *Family of totally umbilical hypersurfaces in an Einstein space*, Ann. of Math. (to appear).

Cambridge, Mass.

ON THE DIRECT PRODUCT OF BANACH SPACES

BY
ROBERT SCHATTEN

Introduction. Two Banach spaces E_1 and E_2 may be combined in two different ways; the well known $E_1 \oplus E_2$ and $E_1 \otimes E_2$. While $E_1 \oplus E_2$ refers to a space of pairs $\{f, \varphi\}$ which are added vectorially, $E_1 \otimes E_2$ is a linear vector space determined by "products" $f \otimes \varphi$, and the $f \otimes \varphi$'s are linearly independent except when we have a relation, which is a consequence of the fact that the $\otimes$ operator is distributive; for instance

$$(f_1 + f_2) \otimes (\varphi_1 + \varphi_2) = f_1 \otimes \varphi_1 + f_1 \otimes \varphi_2 + f_2 \otimes \varphi_1 + f_2 \otimes \varphi_2.$$

The notion of $E_1 \otimes E_2$ for the case of finite dimensions has been already mentioned by H. Weyl [16](1). With each vector e_1 resp. e_2 in a space E_1 of m, resp. E_2 of n dimensions, there is associated a vector $e_1 \otimes e_2$ in the space of $m \cdot n$ dimensions. The totality of vectors $e_1 \otimes e_2$ do not themselves constitute a linear manifold, but their linear combinations fill the entire "product space" $E_1 \otimes E_2$.

The operator $\otimes$ has been used for finite-dimensional l_p (termed $l_{p,n}$) by F. J. Murray, in order to show that there exist linear manifolds without complements [8]. It has also been used by the same author in treating bilinear transformations in Hilbert spaces [9].

The algebraic aspects of the $\otimes$ operator, for the case of finite-dimensional spaces, has been discussed by Hitchcock [5, 6], and Oldenburger [13, 14].

The study of the $\otimes$ operator for infinite-dimensional spaces requires a more abstract method. A complete discussion for Hilbert spaces has been given by F. J. Murray and J. von Neumann [10], who did not make use of the existence of a basis. A few special results for L_p spaces have been obtained by Bourgin [2].

It should be pointed out, however, that so far the study of the $\otimes$ operator, assumed either the existence of an inner product, or a basis, or a projection with bound 1.

The object of this paper is the study of the $\otimes$ operator in a most general form, for any Banach spaces.

For $f \in E_1$, $\varphi \in E_2$, we construct "products" $f \otimes \varphi$. With these we form a linear set $\mathfrak{A}(E_1, E_2)$ consisting of all "expressions" (that is, finite sums) $\sum f_i \otimes \varphi_i$. These expressions must first be considered algebraically, however, since the distributive property introduces certain linear dependencies (§§1, 2). The next problem is that of defining a norm, and the space $E_1 \otimes E_2$ is obtained

Presented to the Society, October 25, 1941; received by the editors January 26, 1942.

(1) Numerals in brackets refer to bibliography at the end of this paper.

by "closing up" the set $\mathfrak{A}(E_1, E_2)$. We are only interested in those norms which are "crossnorms" that is, $\|f\otimes\varphi\| = \|f\|\,\|\varphi\|$ for $f\in E_1$, $\varphi\in E_2$. With the expressions $\sum f_i\otimes\varphi_i$ it is desirable to consider the set $\mathfrak{A}(\overline{E}_1, \overline{E}_2)$ of expressions

$$\sum F_i \otimes \phi_i, \qquad \text{where } F_i \in \overline{E}_1,\ \phi_i \in \overline{E}_2.$$

For a given norm in $\mathfrak{A}(E_1, E_2)$ we construct an associate norm in $\mathfrak{A}(\overline{E}_1, \overline{E}_2)$ (§3). We prove the existence of the greatest crossnorm, the least crossnorm whose associate is also a crossnorm (§4), as well as that of a "self-associate" crossnorm, which is a generalization of the crossnorm given for Hilbert spaces by F. J. Murray and J. von Neumann [10], (§5).

Finally, we mention a few unsolved problems in connection with the work in preceding sections (§6), and indicate possibilities for the construction of "general crossnorms" (§7).

These problems were suggested to me by Professor F. J. Murray, who also pointed out the algebraic discussions given in §§1, 2.

1. Let E_1 and E_2 denote two linear spaces. We introduce two symbols $\otimes$ and $\cdot+\cdot$. With these for $f_1, \cdots, f_n$ in E_1 and $\varphi_1, \cdots, \varphi_n$ in E_2 we construct formal expressions $f_1\otimes\varphi_1\cdot+\cdot f_2\otimes\varphi_2\cdot+\cdot\cdots\cdot+\cdot f_n\otimes\varphi_n$. We may abbreviate this by writing $\sum_{i=1}^n f_i\otimes\varphi_i$. Between these we introduce a relation $\sim$ subject to the following rules:

1. If P is a permutation on $1, 2, \cdots, n$, and $P(i)$ denotes the integer into which P takes i, then

$$\sum_{i=1}^n f_i \otimes \varphi_i \sim \sum_{i=1}^n f_{P(i)} \otimes \varphi_{P(i)}.$$

2a. $(f_1'+f_1'')\otimes\varphi_1\cdot+\cdot f_2\otimes\varphi_2\cdot+\cdot\cdots\cdot+\cdot f_n\otimes\varphi_n\sim f_1'\otimes\varphi_1\cdot+\cdot f_1''\otimes\varphi_1\cdot+\cdot f_2\otimes\varphi_2\cdot+\cdot\cdots\cdot+\cdot f_n\otimes\varphi_n$.

2b. $f_1\otimes(\varphi_1'+\varphi_1'')\cdot+\cdot f_2\otimes\varphi_2\cdot+\cdot\cdots\cdot+\cdot f_n\otimes\varphi_n\sim f_1\otimes\varphi_1'\cdot+\cdot f_1\otimes\varphi_1''\cdot+\cdot f_2\otimes\varphi_2\cdot+\cdot\cdots\cdot+\cdot f_n\otimes\varphi_n$.

3. $(a_1f_1)\otimes\varphi_1\cdot+\cdot(a_2f_2)\otimes\varphi_2\cdot+\cdot\cdots\cdot+\cdot(a_nf_n)\otimes\varphi_n\sim f_1\otimes(a_1\varphi_1)\cdot+\cdot f_2\otimes(a_2\varphi_2)\cdot+\cdot\cdots\cdot+\cdot f_n\otimes(a_n\varphi_n)$.

Definition 1.1. *Two expressions $\sum_{i=1}^n f_i\otimes\varphi_i$ and $\sum_{j=1}^m g_j\otimes\psi_j$ will be termed equivalent, if one can be transformed into the other by a finite number of successive applications of Rules* 1, 2, 3. *We write this $\sum_{i=1}^n f_i\otimes\varphi_i\simeq\sum_{j=1}^m g_j\otimes\psi_j$.*

Rule 1 shows that $\simeq$ is reflexive, that is, every expression is equivalent to itself. The definition also implies transitivity.

A number of elementary results can be easily obtained. For instance, if

$$\sum_{i=1}^n f_i \otimes \varphi_i \simeq \sum_{i=1}^{n'} f_i' \otimes \varphi_i' \quad\text{and}\quad \sum_{j=1}^m g_j \otimes \psi_j \simeq \sum_{j=1}^{m'} g_j' \otimes \psi_j'$$

then

$$\sum_{i=1}^{n} f_i \otimes \varphi_i \cdot + \cdot \sum_{j=1}^{m} g_j \otimes \psi_j \simeq \sum_{i=1}^{n'} f_i' \otimes \varphi_i' \cdot + \cdot \sum_{j=1}^{m'} g_j' \otimes \psi_j' .$$

Because of Rules 2a, 2b, n is not an invariant under equivalence for the expression $\sum_{i=1}^{n} f_i \otimes \varphi_i$. However, we now define a quantity "the rank of $\sum_{i=1}^{n} f_i \otimes \varphi_i$" which will be shown to be an invariant.

DEFINITION 1.2. *Consider $\sum_{i=1}^{n} f_i \otimes \varphi_i$. Let us suppose that the set of f_i's is k-dimensional and that the set of φ_i's is l-dimensional. Let $g_1, \cdots, g_k$ be k linearly independent elements in the set of linear combinations of the f_i, and $\psi_1, \cdots, \psi_l$ be l linearly independent elements in the set of linear combinations of the φ_i. Then*

$$f_i = \sum_{p=1}^{k} a_p^{(i)} g_p, \qquad \varphi_i = \sum_{q=1}^{l} b_q^{(i)} \psi_q \qquad \textit{for } 1 \leqq i \leqq n,$$

and

$$\begin{aligned}\sum_{i=1}^{n} f_i \otimes \varphi_i &\simeq \sum_{i=1}^{n} \left(\sum_{p=1}^{k} a_p^{(i)} g_p \right) \otimes \left(\sum_{q=1}^{l} b_q^{(i)} \psi_q \right) \\ &\simeq \sum_{p=1}^{k} \sum_{q=1}^{l} \left(\sum_{i=1}^{n} a_p^{(i)} b_q^{(i)} \right) g_p \otimes \psi_q \\ &\simeq \sum_{p=1}^{k} \sum_{q=1}^{l} a_{p,q} g_p \otimes \psi_q .\end{aligned}$$

We define the "rank of the expression $\sum_{i=1}^{n} f_i \otimes \varphi_i$" as the rank of the matrix $(a_{p,q})$ $p=1, \cdots, k$; $q=1, \cdots, l$.

To justify this definition we show that the rank does not depend upon the choice of the g_p or ψ_q. For suppose we had taken $g_1', \cdots, g_k'$ and $\psi_1', \cdots, \psi_l'$ instead of $g_1, \cdots, g_k$ and $\psi_1, \cdots, \psi_l$ above. Then we can show that

$$\sum_{i=1}^{n} f_i \otimes \varphi_i \simeq \sum_{p=1}^{k} \sum_{q=1}^{l} a_{p,q}' g_p' \otimes \psi_q'$$

where $(a_{p,q}')$ is equal to a product $B(a_{p,q})C$; B resp. C being nonsingular square matrices with k resp. l rows. But the matrix $B(a_{p,q})C$ is of the same rank as $(a_{p,q})$. Thus the rank of an expression is independent of the choice of the g_p and ψ_q.

Let us define f_q^* by means of the equation $f_q^* = \sum_{p=1}^{k} a_{p,q} g_p$. Then, inasmuch as the ψ_j are linearly independent, the rank of the expression $\sum_{i=1}^{n} f_i \otimes \varphi_i$ is the number of f_q^* which are linearly independent.

LEMMA 1.1. *The rank r of an expression $\sum_{i=1}^{n} f_i \otimes \varphi_i$ is an invariant under equivalence* (*Definition* 1.1).

Proof. It is easily seen that Rule 1 does not affect r. In 2a, let $f_1=f_1'+f_1''$, $h=(1/2)(f_1'-f_1'')$. Then $f_1'=(1/2)f_1+h$, $f_1''=(1/2)f_1-h$,

$$f_1\otimes\varphi_1\cdot+\cdot\sum_{i=2}^n f_i\otimes\varphi_i\simeq((1/2)f_1+h)\otimes\varphi_1$$
$$\cdot+\cdot((1/2)f_1-h)\otimes\varphi_1\cdot+\cdot\sum_{i=2}^n f_i\otimes\varphi_i.$$

Now, if $h=\sum_{p=1}^k a_p g_p$, a calculation shows that the f_q^* in the above are the same for both sides of the relation. On the other hand, if h is linearly independent of $g_1, \cdots, g_k$, we may put $h=g_{k+1}$. But g_{k+1} will appear in the f_q^* only with zero coefficients and again the f_q^* are the same for both sides. Since r is the number of linearly independent f_q^*, r must be also the same on both sides.

A similar discussion will show that r is also unaffected by 2b.

In 3, if any term a_i is zero, we may take f_i as zero. If φ_i is also zero, we may disregard the term. If φ_i is not zero, it may be taken as ψ_1. Then it is easily seen that this term does not contribute to the f_q^*, and hence does not affect the rank. Thus an expression $\sum_{i=1}^n(a_if_i)\otimes\varphi_i$ has the same rank as an expression in which the terms with $a_i=0$ are disregarded. A similar statement holds for $\sum_{i=1}^n f_i\otimes(a_i\varphi_i)$.

This implies that we need consider only the case in which each $a_i\neq 0$. But here we may take the same g_p and ψ_q on both sides of the relation. We then see that $a_p^{(i)}b_q^{(i)}$ has the same value for both sides. Hence the matrix $(a_{p,q})$ and the rank are the same in both cases.

Thus the rank is unaffected by each of these rules. It follows that it is unaffected by any sequence of applications of these rules, and hence it is invariant under equivalence.

LEMMA 1.2. *Every expression $\sum_{i=1}^n f_i\otimes\varphi_i$ is equivalent to either $0\otimes 0$, or to an expression $\sum_{i=1}^m g_i\otimes\psi_i$ in which both the $g_1, \cdots, g_m$ and $\psi_1, \cdots, \psi_m$ are linearly independent. Furthermore, m equals the rank of $\sum_{i=1}^n f_i\otimes\varphi_i$.*

Proof. It is readily seen that if in either of the sets $f_1, \cdots, f_n; \varphi_1, \cdots, \varphi_n$ the elements are linearly dependent, then $\sum_{i=1}^n f_i\otimes\varphi_i$ is equivalent to an expression involving only $n-1$ terms. For instance, if $f_1=\sum_{i=2}^n a_if_i$, then

$$\begin{aligned} f_1\otimes\varphi_1\cdot+\cdot\sum_{i=2}^n f_i\otimes\varphi_i &\simeq\left(\sum_{i=2}^n a_if_i\right)\otimes\varphi_1\cdot+\cdot\sum_{i=2}^n f_i\otimes\varphi_i\\ &\simeq\sum_{i=2}^n(a_if_i)\otimes\varphi_1\cdot+\cdot\sum_{i=2}^n f_i\otimes\varphi_i\\ &\simeq\sum_{i=2}^n f_i\otimes(a_i\varphi_1)\cdot+\cdot\sum_{i=2}^n f_i\otimes\varphi_i\\ &\simeq\sum_{i=2}^n f_i\otimes(a_i\varphi_1+\varphi_i). \end{aligned}$$

We may therefore continue to reduce the number of terms until we have either $\sum_{i=1}^{m} g_i \otimes \psi_i$, in which both the $g_1, \cdots, g_m$ and $\psi_1, \cdots, \psi_m$ are linearly independent or $f \otimes 0$, or $0 \otimes \varphi$. But $f \otimes 0 \simeq f \otimes (0 \cdot 0) \simeq (0f) \otimes 0 \simeq 0 \otimes 0$. Similarly $0 \otimes \varphi \simeq 0 \otimes 0$.

Now the expression $\sum_{i=1}^{m} g_i \otimes \psi_i$ with both the g_i and ψ_i linearly independent has rank m. For the g_i and ψ_i can be used as in Definition 1.2. The resulting matrix is (δ_{ij}), $i, j = 1, \cdots, m$. Since, however, the rank of an expression is invariant under equivalence (Lemma 1.1), m must be also the rank of $\sum_{i=1}^{n} f_i \otimes \varphi_i$.

LEMMA 1.3. *Suppose that in the expression $\sum_{i=1}^{n} f_i \otimes \varphi_i$ the φ_i's are linearly independent. Then the rank of this expression is r, the number of the f_i's which are linearly independent. In particular, if $\sum_{i=1}^{n} f_i \otimes \varphi_i \simeq 0 \otimes 0$, each f_i is zero.*

Proof. If $r = 0$, the rank is zero also. Suppose $r \neq 0$. We may assume that $f_1, \cdots, f_r$ are linearly independent, since otherwise a permutation of the terms (Rule 1) will give this. Then $f_{r+p} = \sum_{k=1}^{r} a_{p,k} f_k$. Hence

$$\sum_{k=1}^{n} f_k \otimes \varphi_k \simeq \sum_{k=1}^{r} f_k \otimes \varphi_k \cdot + \cdot \sum_{p=1}^{n-r} \left(\sum_{k=1}^{r} a_{p,k} f_k \right) \otimes \varphi_{r+p}$$

$$\simeq \sum_{k=1}^{r} f_k \otimes \left(\varphi_k + \sum_{p=1}^{n-r} a_{p,k} \varphi_{r+p} \right) \simeq \sum_{k=1}^{r} f_k \otimes \psi_k$$

where

$$\psi_k = \varphi_k + \sum_{p=1}^{n-r} a_{p,k} \varphi_{r+p}, \qquad k = 1, \cdots, r.$$

Since the φ_i's are linearly independent, that is also true for the ψ_i's. But the rank of $\sum_{k=1}^{r} f_k \otimes \psi_k$ is r (Lemma 1.2), therefore the rank of $\sum_{i=1}^{n} f_i \otimes \varphi_i$ is also r (Lemma 1.1).

Furthermore, from Lemma 1.2, follows that an expression is equivalent to $0 \otimes 0$ if and only if its rank is 0. From the preceding we see that the rank is zero, if and only if each f_i is zero. This implies the last statement of our lemma.

COROLLARY. *If $\sum_{i=1}^{n} f_i \otimes \varphi_i \simeq \sum_{i=1}^{n} g_i \otimes \varphi_i$ and the φ_i are linearly independent, then $f_i = g_i$ for $i = 1, \cdots, n$.*

LEMMA 1.4. *If $\sum_{i=1}^{n} f_i \otimes \varphi_i \simeq 0 \otimes 0$, and k of the f_i and l of the φ_i are linearly independent, then $k + l \leq n$.*

Proof. Suppose that the $\varphi_1, \cdots, \varphi_l$ are linearly independent and, $\varphi_{l+i} = \sum_{j=1}^{l} a_{i,j} \varphi_j$ for $j = 1, \cdots, n-l$. Then

$$0 \otimes 0 \simeq \sum_{j=1}^{n} f_j \otimes \varphi_j \simeq \sum_{j=1}^{l} f_j \otimes \varphi_j \cdot + \cdot \sum_{j=1}^{l} \left(\sum_{i=1}^{n-l} a_{i,j} f_{l+i} \right) \otimes \varphi_j$$

$$\simeq \sum_{j=1}^{l} \left(f_j + \sum_{i=1}^{n-l} a_{i,j} f_{l+i} \right) \otimes \varphi_j.$$

Thus Lemma 1.3 implies $f_j+\sum_{i=1}^{n-l}a_{i,j}f_{l+i}=0$ for $j=1,\cdots,l$. The l relations between the f_i are linearly independent, and hence there can be at most $n-l$ of the f_i linearly independent. Thus $k\leqq n-l$.

LEMMA 1.5. *If* $\sum_{i=1}^{n}f_i\otimes\varphi_i\simeq\sum_{j=1}^{m}g_j\otimes\psi_j$ *and each of the sets* $f_1,\cdots,f_n$; $\varphi_1,\cdots,\varphi_n$; $g_1,\cdots,g_m$; $\psi_1,\cdots,\psi_m$, *are linearly independent, then* $n=m$, *and there exists a square matrix* $(a_{i,j})$, $i,j=1,\cdots,n$, *with an inverse* $(A_{i,j})$, $i,j=1,\cdots,n$, *such that*

$$\psi_k=\sum_{j=1}^{n}a_{k,j}\varphi_j;\qquad g_k=\sum_{j=1}^{n}A_{j,k}f_j.$$

Proof. Lemma 1.3 states that m and n are the ranks of the corresponding expressions. Thus Lemma 1.1 implies that they are equal.

Now $\sum_{i=1}^{n}f_i\otimes\varphi_i\simeq\sum_{j=1}^{n}g_j\otimes\psi_j$ implies

$$\sum_{i=1}^{n}f_i\otimes\varphi_i\cdot+\cdot\sum_{j=1}^{n}(-g_j)\otimes\psi_j\simeq 0\otimes 0. \tag{1}$$

Inasmuch as $f_1,\cdots,f_n$ are linearly independent, Lemma 1.4 implies that at most n of the set $\varphi_1,\cdots,\varphi_n,\psi_1,\cdots,\psi_n$, are linearly independent. Since $\varphi_1,\cdots,\varphi_n$ are linearly independent, each ψ_k must depend upon these. Thus $\psi_k=\sum_{j=1}^{n}a_{k,j}\varphi_j$. But since $\psi_1,\cdots,\psi_n$ are linearly independent, the matrix $(a_{i,j})$, $i,j=1,\cdots,n$ must have an inverse $(A_{i,j})$, $i,j=1,\cdots,n$, and $\varphi_j=\sum_{k=1}^{n}A_{j,k}\psi_k$. Substituting in (1), we get

$$\sum_{k=1}^{n}\left(\sum_{i=1}^{n}A_{i,k}f_i-g_k\right)\otimes\psi_k\simeq 0\otimes 0.$$

Lemma 1.3 now implies that $g_k=\sum_{i=1}^{n}A_{i,k}f_i$.

DEFINITION 1.3. *The set of all expressions of the form* $\sum_{i=1}^{n}f_i\otimes\varphi_i$, *we denote by* $\mathfrak{A}(E_1,E_2)$. *Let* $\bar{f}$ *denote the set of all expressions equivalent to a fixed expression* $\sum_{i=1}^{n}f_i\otimes\varphi_i$. *If an expression is in* $\bar{f}$ *it will be termed "an expression for* $\bar{f}$." *Let* $\mathfrak{A}^*(E_1,E_2)$ *denote the set of such* $\bar{f}$*'s.*

If $\sum_{i=1}^{n}f_i\otimes\varphi_i$ is an expression for $\bar{f}$, $\sum_{j=1}^{m}g_j\otimes\psi_j$ an expression for $\bar{g}$, α a number, we define $\alpha\bar{f}$ as the set of expressions equivalent to $\sum_{i=1}^{n}(\alpha f_i)\otimes\varphi_i$, and $\bar{f}+\bar{g}$ as the set of expressions equivalent to $\sum_{i=1}^{n}f_i\otimes\varphi_i\cdot+\cdot\sum_{j=1}^{m}g_j\otimes\psi_j$.

It is a consequence of Definition 1.1 and Rules 1, 2, 3, that $\alpha\bar{f}$ does not depend on the particular expression used. Similarly $\bar{f}+\bar{g}$ is defined uniquely.

It is easy to see that the usual properties of addition and multiplication by a scalar hold; for instance $\bar{f}+\bar{g}=\bar{g}+\bar{f}$,

$$\bar{f}+(\bar{g}+\bar{h})=(\bar{f}+\bar{g})+\bar{h},\qquad \alpha(\bar{f}+\bar{g})=\alpha\bar{f}+\alpha\bar{g}$$

and

$$(\alpha\beta)\bar{f}=\alpha(\beta\bar{f}).$$

The zero element $\bar{0}$ is the class of all expressions equivalent to $0\otimes 0$. Thus $\mathfrak{A}^*(E_1, E_2)$ is a linear set, that is, a commutative group with scalar operators.

Sometimes we will find it convenient to permit "an expression for $\bar{f}$" to stand for $\bar{f}$.

2. LEMMA 2.1.a. *If F is an additive and homogenous functional on E_1, and $\sum_{i=1}^n f_i\otimes\varphi_i \simeq \sum_{j=1}^m g_j\otimes\psi_j$ then*

$$\sum_{i=1}^n F(f_i)\varphi_i = \sum_{j=1}^m F(g_j)\psi_j.$$

Proof. Consider first $\sum_{i=1}^n f_i\otimes\varphi_i$. Suppose that φ_n is linearly dependent on $\varphi_1, \cdots, \varphi_{n-1}$; $\varphi_n = \sum_{i=1}^{n-1} a_i\varphi_i$. Then $\sum_{i=1}^n f_i\otimes\varphi_i \simeq \sum_{i=1}^{n-1}(f_i + a_i f_n)\otimes\varphi_i$ and

$$\begin{aligned}\sum_{i=1}^n F(f_i)\varphi_i &= \sum_{i=1}^{n-1} F(f_i)\varphi_i + F(f_n)\left(\sum_{i=1}^{n-1} a_i\varphi_i\right)\\ &= \sum_{i=1}^{n-1}(F(f_i) + a_iF(f_n))\varphi_i\\ &= \sum_{i=1}^{n-1} F(f_i + a_if_n)\varphi_i.\end{aligned}$$

A similar statement holds in the case in which f_n is linearly dependent on $f_1, \cdots, f_{n-1}$. These results can be applied successively in such a way that one has finally $\sum_{i=1}^n f_i\otimes\varphi_i \simeq \sum_{i=1}^k f_i' \otimes\varphi_i'$ and $\sum_{i=1}^n F(f_i)\varphi_i = \sum_{i=1}^k F(f_i')\varphi_i'$ with both the f_i''s and φ_i''s linearly independent. Suppose one has gone through the corresponding process with $\sum_{j=1}^m g_j\otimes\psi_j$. The conclusion of our lemma is then a simple consequence of Lemma 1.5.

Let $E_1^\dagger$ resp. $E_2^\dagger$ denote the set of additive and homogenous functionals on E_1 resp. E_2. For $F_1, \cdots, F_n$ in $E_1^\dagger$, and $\phi_1, \cdots, \phi_n$ in $E_2^\dagger$, we construct expressions $\sum_{i=1}^n F_i\otimes\phi_i$.

A similar reasoning to that in Lemma 2.1.a proves

LEMMA 2.1.b. *If $\sum_{i=1}^n F_i\otimes\phi_i \simeq \sum_{j=1}^m G_j\otimes\Psi_j$ then for $f\subset E_1$, we have $\sum_{i=1}^n F_i(f)\phi_i = \sum_{j=1}^m G_j(f)\Psi_j$.*

Combining these results we easily obtain

LEMMA 2.1.c. *If*

$$\sum_{i=1}^n f_i\otimes\varphi_i \simeq \sum_{i=1}^{n'} f_i' \otimes\varphi_i'$$

and

$$\sum_{j=1}^m F_j\otimes\phi_j \simeq \sum_{j=1}^{m'} F_j' \otimes\phi_j'$$

then

$$\left(\sum_{j=1}^{m} F_j \otimes \phi_j\right)\left(\sum_{i=1}^{n} f_i \otimes \varphi_i\right) = \left(\sum_{j=1}^{m'} F_j' \otimes \phi_j'\right)\left(\sum_{i=1}^{n'} f_i' \otimes \varphi_i'\right)$$

where under $(\sum_{j=1}^{m} F_j \otimes \phi_j)(\sum_{i=1}^{n} f_i \otimes \varphi_i)$ *we understand*

$$\sum_{j=1}^{m} \sum_{i=1}^{n} F_j(f_i)\phi_j(\varphi_i).$$

DEFINITION 2.1. *A set S of additive and homogenous functionals on E_1, will be called fundamental if $F(f)=0$ for all $F \in S$, implies $f=0$.*

LEMMA 2.2. *Let S^0 denote a fundamental set contained in $E_1^\dagger$, and S^{00} a fundamental set contained in $E_2^\dagger$. Then if for an expression $\sum_{i=1}^{n} f_i \otimes \varphi_i$, we have $\sum_{i=1}^{n} F(f_i)\phi(\varphi_i)=0$ for all $F \in S^0$, $\phi \in S^{00}$ then $\sum_{i=1}^{n} f_i \otimes \varphi_i \simeq 0 \otimes 0$.*

Proof. Suppose $\sum_{i=1}^{n} f_i \otimes \varphi_i \simeq \sum_{i=1}^{n'} f_i' \otimes \varphi_i'$ where the f_i''s and φ_i''s are linearly independent, then for a certain $F^0 \in S^0$ we have $F^0(f_1') \neq 0$, and since the φ_i''s are linearly independent $\sum_{i=1}^{n'} F^0(f_i')\varphi_i' \neq 0$. Therefore for a certain $\phi^0 \in S^{00}$ we have $\sum_{i=1}^{n'} F^0(f_i')\phi^0(\varphi_i') \neq 0$. Lemma 2.1 gives $\sum_{i=1}^{n} F^0(f_i)\phi^0(\varphi_i) \neq 0$. This completes the proof.

THEOREM 2.1. *Let S^0 resp. S^{00} denote fundamental sets of additive and homogenous functionals on E_1 resp. E_2. Then, a necessary and sufficient condition, for the expressions*

$$\sum_{i=1}^{n} f_i \otimes \varphi_i \quad \textit{and} \quad \sum_{i=1}^{n'} f_i' \otimes \varphi_i'$$

to be equivalent is that

$$\sum_{i=1}^{n} F(f_i)\phi(\varphi_i) = \sum_{i=1}^{n'} F(f_i')\phi(\varphi_i') \qquad \textit{for all } F \in S^0,\ \phi \in S^{00}.$$

Proof. The necessity follows from Lemma 2.1. For the sufficiency put $f_i' = -f_{n+i}$, $\varphi_i' = \varphi_{n+i}$; $1 \leq i \leq n'$, then according to our assumption,

$$\sum_{i=1}^{n+n'} F(f_i)\phi(\varphi_i) = 0 \qquad \text{for } F \in S^0,\ \phi \in S^{00}.$$

Lemma 2.2 gives

$$\sum_{i=1}^{n+n'} f_i \otimes \varphi_i \simeq 0 \otimes 0 \quad \text{or} \quad \sum_{i=1}^{n} f_i \otimes \varphi_i \simeq \sum_{i=1}^{n'} f_i' \otimes \varphi_i'.$$

3. For our further considerations we shall assume that E_1 and E_2 are two Banach spaces, and denote by $\overline{E}_1$ resp. $\overline{E}_2$ the space of linear functionals on E_1 resp. E_2.

DEFINITION 3.1. *Under a norm N in $\mathfrak{A}(E_1, E_2)$ (Definition 1.3) we shall*

understand a non-negative function of expressions satisfying the following conditions:

I. $N(\sum_{i=1}^{n} f_i \otimes \varphi_i) = 0$ *if and only if* $\sum_{i=1}^{n} f_i \otimes \varphi_i \simeq 0 \otimes 0$.

II. $N(\sum_{i=1}^{n} (af_i) \otimes \varphi_i) = |a| N(\sum_{i=1}^{n} f_i \otimes \varphi_i)$ *for any real number* a.

III. $N(\sum_{i=1}^{n} f_i \otimes \varphi_i + \sum_{i=1}^{n'} f_i' \otimes \varphi_i') \leqq N(\sum_{i=1}^{n} f_i \otimes \varphi_i) + N(\sum_{i=1}^{n'} f_i' \otimes \varphi_i')$.

IV. $N(\sum_{i=1}^{n} f_i \otimes \varphi_i) = N(\sum_{i=1}^{n'} f_i' \otimes \varphi_i')$ *if* $\sum_{i=1}^{n} f_i \otimes \varphi_i \simeq \sum_{i=1}^{n'} f_i' \otimes \varphi_i'$.

Definition 3.2. *Consider the set* $\mathfrak{A}(\overline{E}_1, \overline{E}_2)$ *of expressions of the form* $\sum_{j=1}^{m} F_j \otimes \phi_j$, *where* $F_1, \cdots, F_m$ *are in* $\overline{E}_1$, *and* $\phi_1, \cdots, \phi_m$ *are in* $\overline{E}_2$. *For a fixed expression* $\sum_{j=1}^{m_0} F_j^0 \otimes \phi_j^0$ *in* $\mathfrak{A}(\overline{E}_1, \overline{E}_2)$ *we define* $\overline{N}(\sum_{j=1}^{m_0} F_j^0 \otimes \phi_j^0)$ *as the smallest number satisfying the inequality*

$$\left|\left(\sum_{j=1}^{m_0} F_j^0 \otimes \phi_j^0\right)\left(\sum_{i=1}^{n} f_i \otimes \varphi_i\right)\right| \leqq \overline{N}\left(\sum_{j=1}^{m_0} F_j^0 \otimes \phi_j^0\right) N\left(\sum_{i=1}^{n} f_i \otimes \varphi_i\right)$$

for all $\sum_{i=1}^{n} f_i \otimes \varphi_i$ *in* $\mathfrak{A}(E_1, E_2)$.

Thus $\overline{N}$ is a function of expressions in $\mathfrak{A}(\overline{E}_1, \overline{E}_2)$.

Lemma 3.1. *If* $\overline{N}$ *is finite for every expression in* $\mathfrak{A}(\overline{E}_1, \overline{E}_2)$, *then* $\overline{N}$ *also satisfies conditions* I–IV.

Proof. I. If $\sum_{j=1}^{m} F_j \otimes \phi_j \simeq 0 \otimes 0$, then Lemma 2.1.c gives $\overline{N}(\sum_{j=1}^{m} F_j \otimes \phi_j) = 0$. Suppose $\sum_{j=1}^{m} F_j \otimes \phi_j \simeq \sum_{j=1}^{m'} F_j' \otimes \phi_j'$ where the F_j''s and ϕ_j''s are linearly independent (Lemma 2.2). If f in E_1 is such that $F_1'(f) \neq 0$ then $\sum_{j=1}^{m'} F_j'(f)\phi_j' \neq 0$ since the ϕ_j''s are linearly independent. Let $\varphi \in E_2$ be such that $\sum_{j=1}^{m'} F_j'(f)\phi_j'(\varphi) \neq 0$. Lemma 2.1.c gives $(\sum_{j=1}^{m} F_j \otimes \phi_j)(f \otimes \varphi) \neq 0$. This implies $\overline{N}(\sum_{j=1}^{m} F_j \otimes \phi_j) > 0$.

II and III are immediate.

IV is a consequence of Lemma 2.1.c.

For any Banach space we may assume $E \subset \overline{\overline{E}}$.

Lemma 3.2. $\overline{\overline{N}} \leqq N$ *for all expressions in* $\mathfrak{A}(E_1, E_2) \subset \mathfrak{A}(\overline{\overline{E}}_1, \overline{\overline{E}}_2)$.

Proof. Let $f_1, f_2, \cdots$ resp. $\varphi_1, \varphi_2, \cdots$ denote elements in E_1 resp. E_2, $F_1, F_2, \cdots$ resp. $\phi_1, \phi_2, \cdots$, denote elements in $\overline{E}_1$ resp. $\overline{E}_2$; $F_1^{\ddagger}, F_2^{\ddagger}, \cdots$ resp. $\phi_1^{\ddagger}, \phi_2^{\ddagger}, \cdots$ denote elements in $\overline{\overline{E}}_1$ resp. $\overline{\overline{E}}_2$. For an element f^0 in E_1, $F(f^0) = F^{\ddagger 0}(F)$ is an element of $\overline{\overline{E}}_1$, and $\|F^{\ddagger 0}\| = \|f^0\|$. Therefore an expression $\sum_{i=1}^{k} f_i^0 \otimes \varphi_i^0$ will correspond to an expression $\sum_{i=1}^{k} F_i^{\ddagger 0} \otimes \phi_i^{\ddagger 0}$, where the f_i^0 and $F_i^{\ddagger 0}$, as well as the φ_i^0 and $\phi_i^{\ddagger 0}$ are connected by the above mentioned relation. We have

$$\overline{\overline{N}}\left(\sum_{i=1}^{k} f_i^0 \otimes \varphi_i^0\right) = \overline{\overline{N}}\left(\sum_{i=1}^{k} F_i^{\ddagger 0} \otimes \phi_i^{\ddagger 0}\right)$$

$$= \sup\left(\left|\left(\sum_{i=1}^{k} F_i^{\ddagger 0} \otimes \phi_i^{\ddagger 0}\right)\left(\sum_{j=1}^{n} F_j \otimes \phi_j\right)\right| \Big/ \overline{N}\left(\sum_{j=1}^{n} F_j \otimes \phi_j\right)\right)$$

where sup (that is, the least upper bound) is taken over all expressions $\sum_{j=1} F_j \otimes \phi_j$ in $\mathfrak{A}(\overline{E}_1, \overline{E}_2)$, which are not equivalent to $0 \otimes 0$ (of rank 0). Therefore

$$\overline{\overline{N}}\left(\sum_{i=1}^{k} f_i^0 \otimes \phi_i^0\right)$$

$$= \sup\left(\left|\left(\sum_{j=1}^{n} F_j \otimes \phi_j\right)\left(\sum_{i=1}^{k} f_i^0 \otimes \varphi_i^0\right)\right| \Big/ \overline{N}\left(\sum_{j=1}^{n} F_j \otimes \phi_j\right)\right).$$

Let $\epsilon > 0$ be given. There exists an expression $\sum_{j=1}^{n_0} F_j^0 \otimes \phi_j^0$ such that

$$\left|\left(\sum_{j=1}^{n_0} F_j^0 \otimes \phi_j^0\right)\left(\sum_{i=1}^{k} f_i^0 \otimes \varphi_i^0\right)\right| \Big/ \overline{N}\left(\sum_{j=1}^{n_0} F_j^0 \otimes \phi_j^0\right)$$

$$\geqq \overline{\overline{N}}\left(\sum_{i=1}^{k} f_i^0 \otimes \varphi_i^0\right) - \epsilon.$$

From the definition of $\overline{N}$ for a given N, we have

$$\left|\left(\sum_{j=1}^{n_0} F_j^0 \otimes \phi_j^0\right)\left(\sum_{i=1}^{k} f_i^0 \otimes \varphi_i^0\right)\right| \Big/ N\left(\sum_{i=1}^{k} f_i^0 \otimes \varphi_i^0\right) \leqq \overline{N}\left(\sum_{j=1}^{n_0} F_j^0 \otimes \phi_j^0\right).$$

The last two inequalities give

$$N\left(\sum_{i=1}^{k} f_i^0 \otimes \varphi_i^0\right) \geqq \overline{\overline{N}}\left(\sum_{i=1}^{k} f_i^0 \otimes \varphi_i^0\right) - \epsilon.$$

This completes the proof.

LEMMA 3.3. *If N^0 and N^{00} denote two functions in $\mathfrak{A}(E_1, E_2)$ satisfying conditions* I–IV *and $N^0 \leqq N^{00}$, then $\overline{N}^0 \geqq \overline{N}^{00}$.*

The proof is a simple consequence of Definition 3.2 and Lemma 3.1.

Consider the set $\mathfrak{A}^*(E_1, E_2)$ (Definition 1.3). Put $N(\bar{f}) = N(\sum_{i=1}^{n} f_i \otimes \varphi_i)$ for $\sum_{i=1}^{n} f_i \otimes \varphi_i$ in $\bar{f}$. $N(\bar{f})$ is single-valued, as follows from IV for N. Conditions I–III tell us that $N(\bar{f})$ is a norm in $\mathfrak{A}^*(E_1, E_2)$.

Similarly, considering sets of equivalent expressions in $\mathfrak{A}(\overline{E}_1, \overline{E}_2)$ as new elements $\overline{F}$, we obtain a set $\mathfrak{A}^*(\overline{E}_1, \overline{E}_2)$ and a norm $\overline{N}(\overline{F})$ in $\mathfrak{A}^*(\overline{E}_1, \overline{E}_2)$. $\overline{N}(\overline{F})$ will be called the norm associate with N.

We complete $\mathfrak{A}^*(E_1, E_2)$ to $E_1 \otimes E_2$ by adding new elements, namely all fundamental sequences (satisfying Cauchy's condition) of elements in $\mathfrak{A}^*(E_1, E_2)$ with the following conventions:

(α) An element $\bar{f}$ of $\mathfrak{A}^*(E_1, E_2)$ will be considered identical with the sequence $\bar{f}, \bar{f}, \bar{f}, \cdots$.

(β) Two fundamental sequences

$$\bar{f}_1^{(1)}, \bar{f}_2^{(1)}, \bar{f}_3^{(1)}, \cdots; \qquad \bar{f}_1^{(2)}, \bar{f}_2^{(2)}, \bar{f}_3^{(2)}, \cdots$$

are considered identical if and only if

$$\lim_{n\to\infty} N(\bar{f}_n^{(1)} - \bar{f}_n^{(2)}) = 0.$$

(γ) The norm of a fundamental sequence $\bar{f}_1, \bar{f}_2, \bar{f}_3, \cdots$ is by definition $\lim_{n\to\infty} N(\bar{f}_n)$.

In a similar way we complete $\mathfrak{A}^*(\overline{E}_1, \overline{E}_2)$ to $\overline{E}_1 \otimes \overline{E}_2$.

From Lemma 2.1.c follows that $(\sum_{j=1}^m F_j \otimes \phi_j)(\sum_{i=1}^m f_i \otimes \varphi_i)$ depends only upon $\overline{F}$ resp. $\bar{f}$, for which $\sum_{j=1}^m F_j \otimes \phi_j$ resp. $\sum_{i=1}^m f_i \otimes \varphi_i$ are expressions; $\overline{F}(\bar{f})$ is therefore a uniquely defined number for $\bar{f}$ in $\mathfrak{A}^*(E_1, E_2)$, $\overline{F}$ in $\mathfrak{A}^*(\overline{E}_1, \overline{E}_2)$.

Lemma 3.4. *If $\overline{F}_1, \overline{F}_2, \cdots$ resp. $\bar{f}_1, \bar{f}_2, \cdots$ denotes a fundamental sequence of elements in $\mathfrak{A}^*(\overline{E}_1, \overline{E}_2)$ resp. $\mathfrak{A}^*(E_1, E_2)$, then the sequence $\overline{F}_1(\bar{f}_1), \overline{F}_2(\bar{f}_2), \cdots$ is convergent.*

The proof is immediate.

Without fear of misunderstanding, we shall also denote the elements of $E_1 \otimes E_2$ resp. $\overline{E}_1 \otimes \overline{E}_2$ by $\bar{f}$, resp. $\overline{F}$.

From Lemma 3.4 follows that $\overline{F}(\bar{f})$ is uniquely defined for $\bar{f}$ in $E_1 \otimes E_2$, $\overline{F}$ in $\overline{E}_1 \otimes \overline{E}_2$ and $|\overline{F}(\bar{f})| \leqq \overline{N}(\overline{F}) N(\bar{f})$.

Let $\overline{F}^0$ be an element of $\overline{E}_1 \otimes \overline{E}_2$. It is a consequence of Lemma 3.4 that $\overline{F}^0(\bar{f})$ is a linear functional on $E_1 \otimes E_2$. We shall write therefore

$$\overline{E}_1 \otimes \overline{E}_2 \subset \overline{E_1 \otimes E_2}\text{(}^2\text{)}.$$

We shall assume further that N defined on $\mathfrak{A}(E_1, E_2)$ in addition to I–IV satisfies also the following condition of continuity:

V. *$N(\sum_{i=1}^n f_i \otimes \varphi_i)$ is a continuous function of the f_i and φ_i, that is, if $\epsilon > 0$ is given, we can find a $\delta = \delta(f_1, \cdots, f_n; \varphi_1, \cdots, \varphi_n) > 0$, such that for $\|f_i - f_i'\| < \delta$, $\|\varphi_i - \varphi_i'\| < \delta$ for $i = 1, \cdots, n$, we have*

$$N\left(\sum_{i=1}^n f_i \otimes \varphi_i - \sum_{i=1}^n f_i' \otimes \varphi_i'\right) < \epsilon.$$

Lemma 3.5. *If E_1, E_2 are separable and N satisfies conditions* I–V, *then $E_1 \otimes E_2$ is separable.*

Proof. Let $f_1^0, f_2^0, \cdots$ resp. $\varphi_1^0, \varphi_2^0, \cdots$ denote two sequences dense in E_1 resp. E_2. Then the set of expressions $\sum_{i=1}^n f_{t_i}^0 \otimes \varphi_{t_i'}^0$, $t_i, t_i' = 1, 2, 3, \cdots$; $n = 1, 2, 3, \cdots$, is dense in $\mathfrak{A}(E_1, E_2)$. Hence the set of elements $\bar{f}$ for which these expressions stand is dense in $\mathfrak{A}^*(E_1, E_2)$ therefore also in $E_1 \otimes E_2$.

Definition 3.3. *A function N of expressions in $\mathfrak{A}(E_1, E_2)$ is called a cross-norm if in addition to* I–IV *it satisfies the following "cross-property":*

$$N(f \otimes \varphi) = \|f\| \|\varphi\| \qquad \textit{for } f \in E_1, \varphi \in E_2.$$

(2) A supplementary remark is made in Part A, §6.

LEMMA 3.6. *A crossnorm N satisfies condition* V.

Proof. This is an immediate consequence of the following relation

$$N\left(\sum_{i=1}^{n} f_i \otimes \varphi_i - \sum_{i=1}^{n} f_i' \otimes \varphi_i'\right) \leqq N\left(\sum_{i=1}^{n} (f_i - f_i') \otimes \varphi_i\right)$$
$$+ N\left(\sum_{i=1}^{n} f_i \otimes (\varphi_i - \varphi_i')\right) + N\left(\sum_{i=1}^{n} (f_i - f_i') \otimes (\varphi_i - \varphi_i')\right)$$
$$\leqq \sum_{i=1}^{n} \|f_i - f_i'\| \, \|\varphi_i\| + \sum_{i=1}^{n} \|f_i\| \, \|\varphi_i - \varphi_i'\| + \sum_{i=1}^{n} \|f_i - f_i'\| \, \|\varphi_i - \varphi_i'\|.$$

Sometimes we shall assume that the norm N satisfies the following condition:

VI. $\|\sum_{i=1}^{n} F(f_i)\varphi_i\| \leqq \|F\| N(\sum_{i=1}^{n} f_i \otimes \varphi_i)$ *for* F *in* $\overline{E}_1$, *and* $\sum_{i=1}^{n} f_i \otimes \varphi_i$ *in* $\mathfrak{A}(E_1, E_2)$.

DEFINITION 3.4. *From Lemma* 2.1 *follows that* $\sum_{i=1}^{n} F(f_i)\varphi_i$ *is invariant under equivalence. Let* $\bar{f}$ *be an element of* $\mathfrak{A}^*(E_1, E_2)$ *for which* $\sum_{i=1}^{n} f_i \otimes \varphi_i$ *is an expression. We define* $T_{\bar{f}}F$ *as the transformation from* $\overline{E}_1$ *to* E_2 *such that* $T_{\bar{f}}F = \sum_{i=1}^{n} F(f_i)\varphi_i$.

LEMMA 3.7. *If* $T_{\bar{f}} = 0$ *then* $\bar{f} = 0$.

Proof. Let $\bar{f} \neq 0$ and $\sum_{i=1}^{n} f_i \otimes \varphi_i$ be an expression for $\bar{f}$ for which the f_i's and φ_i's are linearly independent. Thus $f_1 \neq 0$. We can find an $F \subset \overline{E}_1$, such that $F(f_1) \neq 0$. This implies $\sum_{i=1}^{n} F(f_i)\varphi_i \neq 0$, since the φ_i's are linearly independent. Thus $T_{\bar{f}}F \neq 0$ and $T_{\bar{f}} \neq 0$. We have shown that $\bar{f} \neq 0$ implies $T_{\bar{f}} \neq 0$. This completes the proof.

LEMMA 3.8. *Condition* II, IV, *and* VI *for* N *imply* I.

Proof. From II and IV follows that $\sum_{i=1}^{n} f_i \otimes \varphi_i \simeq 0 \otimes 0$ implies $N(\sum_{i=1}^{n} f_i \otimes \varphi_i) = 0$. Now let $N(\sum_{i=1}^{n} f_i \otimes \varphi_i) = 0$. VI implies $\|\sum_{i=1}^{n} F(f_i)\varphi_i\| = 0$ for F in $\overline{E}_1$. Hence $\sum_{i=1}^{n} f_i \otimes \varphi_i \simeq 0 \otimes 0$ by Lemma 3.7.

4. Among the functions N, a particular one which we shall denote by N_1 will be of great interest to us.

DEFINITION 4.1. *Let* $\sum_{i=1}^{m} f_i^0 \otimes \varphi_i^0$ *be a fixed expression. We define* $N_1(\sum_{i=1}^{n} f_i^0 \otimes \varphi_i^0) = \inf \sum_{i=1}^{n} \|f_i\| \, \|\varphi_i\|$ *where* inf (*the greatest lower bound*) *is taken over the set of all expressions* $\sum_{i=1}^{n} f_i \otimes \varphi_i$ *equivalent to* $\sum_{i=1}^{m} f_i^0 \otimes \varphi_i^0$.

LEMMA 4.1. N_1 *satisfies conditions* II, III, IV, *and* VI, *therefore also* I by *Lemma* 3.8; N_1 *is a crossnorm, therefore it satisfies* V *as follows from Lemma* 3.6.

Proof. That N_1 satisfies II and IV is obvious We proceed to prove III. Let $\sum_{i=1}^{m} f_i^0 \otimes \varphi_i^0$ and $\sum_{i=1}^{n} g_i^0 \otimes \psi_i^0$ be two given expressions, and let $\epsilon > 0$ be given.

Take an expression

$$\sum_{i=1}^{r} f_i \otimes \varphi_i \simeq \sum_{i=1}^{m} f_i^0 \otimes \varphi_i^0$$

such that

$$\sum_{i=1}^{r} \|f_i\| \|\varphi_i\| \leqq N_1\left(\sum_{i=1}^{m} f_i^0 \otimes \varphi_i^0\right) + \epsilon/2.$$

Similarly we find

$$\sum_{i=1}^{s} g_i \otimes \psi_i \simeq \sum_{i=1}^{n} g_i^0 \otimes \psi_i^0$$

such that

$$\sum_{i=1}^{s} \|g_i\| \|\psi_i\| \leqq N_1\left(\sum_{i=1}^{n} g_i^0 \otimes \psi_i^0\right) + \epsilon/2.$$

We have

$$\sum_{i=1}^{r} f_i \otimes \varphi_i \dot{+} \sum_{i=1}^{s} g_i \otimes \psi_i \simeq \sum_{i=1}^{m} f_i^0 \otimes \varphi_i^0 \dot{+} \sum_{i=1}^{n} g_i^0 \otimes \psi_i^0.$$

Condition IV and Definition 4.1 give

$$\begin{aligned} N_1\left(\sum_{i=1}^{m} f_i^0 \otimes \varphi_i^0 \dot{+} \sum_{i=1}^{n} g_i^0 \otimes \psi_i^0\right) &\leqq \sum_{i=1}^{r} \|f_i\| \|\varphi_i\| + \sum_{i=1}^{s} \|g_i\| \|\psi_i\| \\ &\leqq N_1\left(\sum_{i=1}^{m} f_i^0 \otimes \varphi_i^0\right) + N_1\left(\sum_{i=1}^{n} g_i^0 \otimes \psi_i^0\right) + \epsilon. \end{aligned}$$

This proves III.

We proceed to prove VI. Let $\sum_{i=1}^{n} f_i \otimes \varphi_i$ be an expression for $\bar{f}$. We have (Definition 3.4)

$$\|T_{\bar{f}}F\| = \left\|\sum_{i=1}^{n} F(f_i)\varphi_i\right\| \leqq \|F\| \sum_{i=1}^{n} \|f_i\| \|\varphi_i\|.$$

This holds for every expression in $\bar{f}$. Hence

$$\|T_{\bar{f}}F\| \leqq \|F\| \cdot \inf \sum_{i=1}^{n} \|f_i\| \|\varphi_i\| = \|F\| N_1(\bar{f}).$$

To prove the cross-property, we assume $f \neq 0$ and $\varphi \neq 0$ (for $f = 0$ or $\varphi = 0$, the proof is obvious). Let $F \in \overline{E}_1$ be such that $F(f) = \|f\|$, $\|F\| = 1$. If $\sum_{i=1}^{n} f_i \otimes \varphi_i \simeq f \otimes \varphi$, we have (Lemma 2.1)

$$\|F(f)\varphi\| = \left\|\sum_{i=1}^{n} F(f_i)\varphi_i\right\| \quad \text{or} \quad \|f\| \|\varphi\| \leqq \sum_{i=1}^{n} \|f_i\| \|\varphi_i\|.$$

Therefore $N_1(f \otimes \varphi) = \|f\| \|\varphi\|$. This completes the proof.

LEMMA 4.2. *N_1 is the greatest crossnorm.*

Proof. For any crossnorm N, we have

$$N\left(\sum_{i=1}^{k} \overset{0}{f_i} \otimes \overset{0}{\varphi_i}\right) \leqq \sum_{i=1}^{k} N(\overset{0}{f_i} \otimes \overset{0}{\varphi_i}) = \sum_{i=1}^{k} \|\overset{0}{f_i}\| \, \|\overset{0}{\varphi_i}\|.$$

Condition IV for N implies

$$N\left(\sum_{i=1}^{k} \overset{0}{f_i} \otimes \overset{0}{\varphi_i}\right) \leqq \inf \sum_{i=1}^{n} \|f_i\| \, \|\varphi_i\| = N_1\left(\sum_{i=1}^{k} \overset{0}{f_i} \otimes \overset{0}{\varphi_i}\right)$$

where inf is taken over the set of all expressions $\sum_{i=1}^{n} f_i \otimes \varphi_i$ equivalent to $\sum_{i=1}^{k} \overset{0}{f_i} \otimes \overset{0}{\varphi_i}$.

Lemma 4.3. *The norm $\overline{N}$ associated with a crossnorm N, satisfies the condition $\overline{N}(F \otimes \phi) \geqq \|F\| \, \|\phi\|$ for $F \subset \overline{E}_1$, $\phi \subset \overline{E}_2$.*

The proof is immediate.

Definition 4.2. *We define N_0:*

$$N_0\left(\sum_{i=1}^{n} f_i \otimes \varphi_i\right) = \sup\left(\left|\sum_{i=1}^{n} F(f_i)\phi(\varphi_i)\right| \Big/ \|F\| \, \|\phi\|\right)$$

where sup (*the least upper bound*) *is taken over the set of numbers obtained when F resp. ϕ varies in $\overline{E}_1$ resp. $\overline{E}_2$.*

Lemma 4.4. *N_0 and $\overline{N}_0$ are crossnorms.*

Proof. It is not difficult to verify that N_0 is a crossnorm. We shall prove that $\overline{N}_0$ is a crossnorm. We have

$$\begin{aligned}
&\left(\left|(F^0 \otimes \phi^0)\left(\sum_{i=1}^{n} f_i \otimes \varphi_i\right)\right| \Big/ N_0\left(\sum_{i=1}^{n} f_i \otimes \varphi_i\right)\right) \\
&\qquad \leqq \left|(F^0 \otimes \phi^0)\left(\sum_{i=1}^{n} f_i \otimes \varphi_i\right)\right| \left(\|F^0\| \, \|\phi^0\| \Big/ \left|(F^0 \otimes \phi^0)\left(\sum_{i=1}^{n} f_i \otimes \varphi_i\right)\right|\right) \\
&\qquad = \|F^0\| \, \|\phi^0\|.
\end{aligned}$$

Hence (Definition 3.2) $\overline{N}_0 (F^0 \otimes \phi^0) \leqq \|F^0\| \, \|\phi^0\|$. This together with Lemma 4.3 concludes the proof.

Lemma 4.5. *The associate $\overline{N}$ with a crossnorm $N \geqq N_0$ is also a crossnorm.*

Proof. Lemma 3.3 gives $\overline{N} \leqq \overline{N}_0$. In particular $\overline{N}(F \otimes \phi) \leqq \overline{N}_0(F \otimes \phi) = \|F\| \, \|\phi\|$ by Lemma 4.4. An application of Lemma 4.3 concludes the proof.

Theorem 4.1. *The associate $\overline{N}$ with a crossnorm N is also a crossnorm if and only if $N \geqq N_0$; N_0 is therefore the least crossnorm whose associate is also a crossnorm.*

Proof. The sufficiency is proved in Lemma 4.5. We shall prove the necessity. Suppose that for a crossnorm N and a certain expression $\sum_{i=1}^{n} f_i^0 \otimes \varphi_i^0$ we have

$$N\left(\sum_{i=1}^{n} f_i^0 \otimes \varphi_i^0\right) < N_0\left(\sum_{i=1}^{n} f_i^0 \otimes \varphi_i^0\right).$$

Then, there exists an $F^0 \in \overline{E}_1$ and $\phi^0 \in \overline{E}_2$ such that

$$N\left(\sum_{i=1}^{n} f_i^0 \otimes \varphi_i^0\right) < \left|\sum_{i=1}^{n} F^0(f_i^0)\phi^0(\varphi_i^0)\right| \Big/ \|F^0\| \, \|\phi^0\|$$

and consequently

$$\overline{N}(F^0 \otimes \phi^0) \geqq \left|(F^0 \otimes \phi^0)\left(\sum_{i=1}^{n} f_i^0 \otimes \varphi_i^0\right)\right| \Big/ N\left(\sum_{i=1}^{n} f_i^0 \otimes \phi_i^0\right) > \|F^0\| \, \|\phi^0\|$$

or $\overline{N}$ is not a crossnorm. This completes the proof([3]).

Theorem 4.1.1. *A crossnorm N satisfies condition* VI *if and only if its associate $\overline{N}$ is also a crossnorm.*

The proof is similar to that in Theorem 4.1.

In our future work, speaking about "the least crossnorm," we shall have in mind "the least crossnorm whose associate is also a crossnorm," namely N_0.

5. The least crossnorm N_0, as well as the greatest crossnorm N_1, are defined for any two Banach spaces E_1, E_2. For this reason we shall call them general crossnorms. Similarly $(1/2)(N_0+N_1)$, $\overline{(1/2)(N_0+N_1)}$, $\cdots$ are general crossnorms([4]).

Let K denote the smallest class of crossnorms satisfying the following conditions:

1. $N_1 \in K$.
2. If $N \in K$, then $\overline{N} \in K$.
3. If N^0 and N^{00} belong to K, then $aN^0 + (1-a)N^{00}$ belongs to K for $0 < a < 1$.
4. If $N^0, N^{00}, N^{000}, \cdots$ denotes a monotonic sequence of crossnorms belonging to K, then its limit (which exists because the sequence is bounded by the least and greatest crossnorms) also is in K.

Definition 5.1. *Crossnorms in K are defined for any two Banach spaces E_1, E_2. For this reason we shall call them general crossnorms.*

Lemma 5.1. *If $N^{(1)}$ and $N^{(2)}$ denote two crossnorms* (*not necessarily general*) *and a, b real numbers, such that $a+b=1$, $0<a<1$, then*

([3]) An immediate problem is mentioned in Part B, §6.

([4]) For a general crossnorm N, $\overline{N}$ is to have the following significance in $\mathfrak{A}^*(E_1, E_2)$. We consider N on $\mathfrak{A}^*(\overline{E}_1, \overline{E}_2)$. For this there is an $\overline{N}$ defined on $\mathfrak{A}^*(\overline{\overline{E}}_1, \overline{\overline{E}}_2)$. We consider the latter confined to $\mathfrak{A}^*(E_1, E_2) \subseteq \mathfrak{A}^*(\overline{\overline{E}}_1, \overline{\overline{E}}_2)$.

$$\overline{a\, N^{(1)} + b\, N^{(2)}} \leqq a\, \overline{N}^{(1)} + b\, \overline{N}^{(2)}.$$

Proof. We shall prove that for any $\overline{F}_0 \in \mathfrak{A}^*(\overline{E}_1, \overline{E}_2)$ we have

$$\sup \frac{|\overline{F}_0(\bar{f})|}{aN^{(1)}(\bar{f}) + bN^{(2)}(\bar{f})} \leqq a \sup \frac{|\overline{F}_0(\bar{f})|}{N^{(1)}(\bar{f})} + b \sup \frac{|\overline{F}_0(\bar{f})|}{N^{(2)}(\bar{f})}$$

where sup is taken over all $\bar{f}$'s ($\neq 0$) in $\mathfrak{A}^*(E_1, E_2)$. Suppose that the contrary holds, that is, for a certain $\overline{F}_0$ the last inequality does not hold. Then for a certain $\bar{f}_0$ in $\mathfrak{A}^*(E_1, E_2)$ we have

$$\frac{|\overline{F}_0(\bar{f}_0)|}{a\, N^{(1)}(\bar{f}_0) + bN^{(2)}(\bar{f}_0)} > a \frac{|\overline{F}_0(\bar{f}_0)|}{N^{(1)}(\bar{f}_0)} + b \frac{|\overline{F}_0(\bar{f}_0)|}{N^{(2)}(\bar{f}_0)}.$$

This gives: $\{N^{(1)}(\bar{f}_0) - N^{(2)}(\bar{f}_0)\}^2 < 0$. This cannot happen, and therefore the proof is completed.

Corollary. *Let N^0 and N^{00} denote two general crossnorms, and a, b real numbers, such that $a+b=1$, $0<a<1$, then*

$$\overline{aN^0 + bN^{00}} \leqq a\, \overline{N}^0 + b\, \overline{N}^{00}.$$

Lemma 5.2. *If N denotes a general crossnorm and $\overline{N}$ its associate, then putting $N^{(a)} = aN + (1-a)\overline{N}$, $0 \leqq a \leqq 1$, we have $N^{(a)} + \overline{N}^{(a)} \leqq N + \overline{N}$*(6).

Proof. For $a=1$ the statement is trivial, For $a=0$, $\overline{N} + \overline{\overline{N}} \leqq \overline{N} + N$ as follows from Lemma 3.2, For $0<a<1$, Lemma 5.1 gives $N^{(a)} + \overline{N}^{(a)} \leqq \{aN + (1-a)\overline{N}\} + \{a\overline{N} + (1-a)\overline{\overline{N}}\} \leqq N + \overline{N}$ as follows from Lemma 3.2.

Corollary. *For a general crossnorm N, we have $\overline{(1/2)(N+\overline{N})} \leqq (1/2)(N+\overline{N})$.*

Proof. $N^{(1/2)} + \overline{N}^{(1/2)} \leqq N + \overline{N} = 2N^{(1/2)}$ (Lemma 5.2). Hence $\overline{N}^{(1/2)} \leqq N^{(1/2)}$. This completes the proof.

We shall show an immediate application of the last result.

Theorem 5.1. *There exists a general crossnorm N, with the following properties:*

α. N is identical with its associate $\overline{N}$.

β. $N = \lim_{n\to\infty} N^{(n)}$, where $\{N^{(n)}\}$ is a monotonic sequence of general crossnorms defined in the following way: Let N_1 denote the greatest crossnorm. Put $N^{(1)} = (1/2)(N_1 + \overline{N}_1)$. Let $n>1$, and suppose we have defined $N^{(k)}$ for $k<n$, then put $N^{(n)} = (1/2)(N^{(n-1)} + \overline{N}^{(n-1)})$.

Proof. Since N_1 is the greatest crossnorm, its associate $\overline{N}_1$ is also a crossnorm (Lemma 4.5), and therefore $\overline{N}_1 \leqq N_1$. Lemma 3.3 and the corollary to Lemma 5.2 give

(6) For simplicity of notation $\overline{N}^{(a)}$ shall denote the associate with $N^{(a)}$.

$$\overline{N}_1 \leqq \overline{N}^{(1)} \leqq N^{(1)} \leqq N_1.$$

The last inequality proves that $\overline{N}^{(1)}$ is a crossnorm, therefore $N^{(2)}$ is also a crossnorm. Similarly we obtain

$$\overline{N}_1 \leqq \overline{N}^{(1)} \leqq \overline{N}^{(2)} \leqq N^{(2)} \leqq N^{(1)} \leqq N_1.$$

This proves that $\overline{N}^{(2)}$ is also a crossnorm, therefore $N^{(3)}$ is a crossnorm. Repeating the same process, we obtain a decreasing sequence of crossnorms

$$N_1 \geqq N^{(1)} \geqq N^{(2)} \geqq \cdots$$

and an increasing sequence of its associate crossnorms

$$\overline{N}_1 \leqq \overline{N}^{(1)} \leqq \overline{N}^{(2)} \leqq \cdots .$$

The first sequence is bounded from below by $\overline{N}_1$, hence is convergent; let N denote its limit. The second sequence is bounded from above by N_1, hence convergent; let $\mathfrak{N}$ denote its limit. Therefore:

(1) $\lim_{n\to\infty} N^{(n)} = N$; $\lim_{n\to\infty} \overline{N}^{(n)} = \mathfrak{N}$.

It is easy to verify that

$$N^{(n)} - \overline{N}^{(n)} \leqq (1/2^n)(N_1 - \overline{N}_1) \qquad \text{for } n = 1, 2, \cdots$$

because

$$\begin{aligned} N^{(1)} - \overline{N}^{(1)} &= (1/2)(N_1 + \overline{N}_1) - \overline{N}^{(1)} \leqq (1/2)(N_1 + \overline{N}_1) - \overline{N}_1 \\ &= (1/2)(N_1 - \overline{N}_1), \\ N^{(2)} - \overline{N}^{(2)} &= (1/2)(N^{(1)} + \overline{N}^{(1)}) - \overline{N}^{(2)} \leqq (1/2)(N^{(1)} + \overline{N}^{(1)}) - \overline{N}^{(1)} \\ &= (1/2)(N^{(1)} - \overline{N}^{(1)}) \leqq (1/2^2)(N_1 - \overline{N}_1), \end{aligned}$$

. . . .

Therefore: $N = \mathfrak{N}$. We shall prove $N = \overline{N}$. Since $N \leqq N^{(n)}$ for $n = 1, 2, \cdots$ $\overline{N} \geqq \overline{N}^{(n)}$ hence (1) gives $\overline{N} \geqq \mathfrak{N}$ or

(2) $\overline{N} \geqq N$.

On the other hand,

$$\overline{N}^{(n)} \leqq \mathfrak{N} = N \qquad \text{for } n = 1, 2, \cdots ;$$

hence $\overline{\overline{N}}^{(n)} \geqq \overline{N}$ and therefore (Lemma 3.2) $N^{(n)} \geqq \overline{N}$. (1) gives

(3) $N \geqq \overline{N}$.

(2) and (3) give the required result(⁶).

THEOREM 5.1.1. *Let N^0 denote a general crossnorm (Definition 5.1). The same construction (which we have applied to N_1 in Theorem 5.1) applied to N^0, will always lead to a crossnorm N satisfying conditions α, β mentioned in Theorem 5.1.*

---

[6] The immediate problem is mentioned in Part C, §6.

Proof. Let us notice that for a general crossnorm N^0, we have always $\overline{N}_1 \leqq N^0 \leqq N_1$, hence also $\overline{N}_1 \leqq \overline{N}^0 \leqq N_1$ by virtue of Lemmas 3.2 and 3.3. This assures that the associate with a general crossnorm is also a crossnorm. We put $N^{(1)} = (1/2)(N^0 + \overline{N}^0)$, $N^{(1)}$ is a crossnorm and $\overline{N}^{(1)} \leqq N^{(1)}$ (Lemma 5.2) is also a crossnorm. We have therefore the situation mentioned in Theorem 5.1. Let $n > 1$, and suppose we have defined $N^{(k)}$ for $k < n$, then put $N^{(n)} = (1/2)(N^{(n-1)} + \overline{N}^{(n-1)})$. As in Theorem 5.1 we prove that $\lim_{n\to\infty} N^{(n)}$ exists; call it N, and $\lim_{n\to\infty} \overline{N}^{(n)}$ exists; call it $\mathfrak{N}$. Further,

$$N^{(n)} - \overline{N}^{(n)} \leqq (1/2^{n-1})(N^{(1)} - \overline{N}^{(1)}) \quad \text{for } n = 1, 2, \cdots;$$

hence $N = \mathfrak{N}$. Finally we prove (in exactly the same way as in Theorem 5.1) that $\overline{N} = N$.

We shall apply our results to Hilbert spaces. Let E_1, E_2, denote two Hilbert spaces, and N a crossnorm in $\mathfrak{A}(E_1, E_2)$ (Definition 1.3).

DEFINITION 5.2. *N will be termed self-associate if for every expression $\sum_{j=1}^{m} F_j \otimes \phi_j$ in $\mathfrak{A}(\overline{E}_1, \overline{E}_2)$, we have*

$$\overline{N}\Big(\sum_{j=1}^{m} F_j \otimes \phi_j\Big) = N\Big(\sum_{j=1}^{m} f_j \otimes \varphi_j\Big),$$

where f_j resp. φ_j is the element in E_1 resp. E_2 for which $F_j(f) \equiv (f, f_j)$ in E_1 resp. $\phi_j(\varphi) \equiv (\varphi, \varphi_j)$ in E_2.

For two Hilbert spaces E_1, E_2, F. J. Murray and J. von Neumann introduce the following crossnorm in $\mathfrak{A}(E_1, E_2)$ [10]:

$$N\left(\sum_{j=1}^{n} f_j \otimes \varphi_j\right) = \left\{\sum_{i=1}^{n}\sum_{k=1}^{n} (f_i, f_k)(\varphi_i, \varphi_k)\right\}^{1/2}$$

where the symbol (f_i, f_k) denotes the inner product. Let $\overline{N}$ denote the associate with N, and let

$$\sum_{j=1}^{m} F_j^0 \otimes \phi_j^0$$

be a fixed expression in $\mathfrak{A}(\overline{E}_1, \overline{E}_2)$. We have

$$\begin{aligned} &\overline{N}\left(\sum_{j=1}^{m} F_j^0 \otimes \phi_j^0\right) \\ (1) \qquad &= \sup\left|\left(\sum_{j=1}^{m} F_j^0 \otimes \phi_j^0\right)\left(\sum_{i=1}^{n} f_i \otimes \varphi_i\right)\right| \Big/ N\left(\sum_{i=1}^{n} f_i \otimes \varphi_i\right) \\ &= \sup\left|\left(\sum_{j=1}^{m}\sum_{i=1}^{n} (f_i, f_j^0)(\varphi_i, \varphi_j^0)\right)\right| \Big/ \left\{\sum_{i=1}^{n}\sum_{k=1}^{n} (f_i, f_k)(\varphi_i, \varphi_k)\right\}^{1/2} \end{aligned}$$

where $F_j^0(f) \equiv (f, f_j^0)$, $\phi_j^0(\varphi) \equiv (\varphi, \varphi_j^0)$. Applying Schwarz's inequality to the numerator, we get

$$\overline{N}\left(\sum_{j=1}^{m} F_j^0 \otimes \phi_j^0\right)$$

$$\leqq \sup\left(\left\{\sum_{j=1}^{m}\sum_{l=1}^{m}(f_j^0, f_l^0)(\varphi_j^0, \varphi_l^0)\right\}^{1/2}\left\{\sum_{i=1}^{n}\sum_{k=1}^{n}(f_i, f_k)(\varphi_i, \varphi_k)\right\}^{1/2}\Big/\left\{\sum_{i=1}^{n}\sum_{k=1}^{n}(f_i, f_k)(\varphi_i, \varphi_k)\right\}^{1/2}\right)$$

or

$$\overline{N}\left(\sum_{j=1}^{m} F_j^0 \otimes \phi_j^0\right) \leqq N\left(\sum_{j=1}^{m} f_j^0 \otimes \varphi_j^0\right). \tag{2}$$

On the other hand, taking in particular $\sum_{j=1}^{m} f_j^0\otimes\varphi_j^0$ for the variable expression $\sum_{i=1}^{n} f_i\otimes\varphi_i$, we get from (1)

$$\overline{N}\left(\sum_{j=1}^{m} F_j^0 \otimes \phi_j^0\right) \geqq \left|\sum_{j=1}^{m}\sum_{l=1}^{m}(f_j^0, f_l^0)(\varphi_j^0, \varphi_l^0)\right|\Big/\left\{\sum_{j=1}^{m}\sum_{l=1}^{m}(f_j^0, f_l^0)(\varphi_j^0, \varphi_l^0)\right\}^{1/2}$$

$$= \left\{\sum_{j=1}^{m}\sum_{l=1}^{m}(f_j^0, f_l^0)(\varphi_j^0, \varphi_l^0)\right\}^{1/2}.$$

This means $\overline{N}(\sum_{j=1}^{m} F_j^0\otimes\phi_j^0) \geqq N(\sum_{j=1}^{m} f_j^0\otimes\varphi_j^0)$. The last inequality together with (2) proves that

$$\overline{N}\left(\sum_{j=1}^{m} F_j^0 \otimes \phi_j^0\right) = N\left(\sum_{j=1}^{m} f_j^0 \otimes \varphi_j^0\right)$$

or the crossnorm introduced by F. J. Murray and J. von Neumann is self-associate in the sense of Definition 5.2.

We shall denote this crossnorm by $S_{M,N}$.

THEOREM 5.2. *If E_1 and E_2 denote two Hilbert spaces, then every self-associate crossnorm in $\mathfrak{A}(E_1, E_2)$ is identical with $S_{M,N}$.*

Proof. Let N be a crossnorm in $\mathfrak{A}(E_1, E_2)$. Since $\overline{E}_1 = E_1$, $\overline{E}_2 = E_2$, for an expression $\sum_{j=1}^{m} F_j^0\otimes\phi_j^0$ in $\mathfrak{A}(\overline{E}_1, \overline{E}_2)$ we have

$$\overline{N}\left(\sum_{j=1}^{m} F_j^0 \otimes \phi_j^0\right) = \overline{N}\left(\sum_{j=1}^{m} f_j^0 \otimes \varphi_j^0\right)$$

$$= \sup\left(\left|\sum_{j=1}^{m}\sum_{i=1}^{n}(f_i, f_j^0)(\varphi_i, \varphi_j^0)\right|\Big/ N\left(\sum_{i=1}^{n} f_i \otimes \varphi_i\right)\right)$$

where $F_j^0(f) \equiv (f, f_j^0)$ in E_1, $\phi_j^0(\varphi) \equiv (\varphi, \varphi_j^0)$ in E_2. Taking in particular $\sum_{j=1}^{m} f_j^0\otimes\varphi_j^0$ for the variable expression $\sum_{j=1}^{n} f_i\otimes\varphi_i$ we get

$$\overline{N}\left(\sum_{j=1}^{m} f_j^0 \otimes \varphi_j^0\right) N\left(\sum_{j=1}^{m} f_j^0 \otimes \varphi_j^0\right) \geqq \left\{S_{M,N}\left(\sum_{j=1}^{m} f_j^0 \otimes \varphi_j^0\right)\right\}^2.$$

In particular if N is self-associate in the sense of Definition 5.2, that is, $\overline{N}(\sum_{j=1}^{m} f_j^0 \otimes \varphi_j^0) = N(\sum_{j=1}^{m} f_j^0 \otimes \varphi_j^0)$ we get

$$N^2 \geqq S_{M,N}^2 \quad \text{or} \quad N \geqq S_{M,N}.$$

Taking the associate crossnorms for both sides of the last inequality, we get $\overline{N} \leqq \overline{S}_{M,N}$, or since N and $S_{M,N}$ are self-associate $N \leqq S_{M,N}$. This gives $N = S_{M,N}$, and the proof is completed.

THEOREM 5.3. *If E_1 and E_2 denote two Hilbert spaces, then the general crossnorm N constructed in Theorem 5.1 by means of the greatest crossnorm N_1 coincides with the usual self-associate crossnorm $S_{M,N}$.*

Proof. From the construction of N, and $\overline{E}_1 = E_1$, $\overline{E}_2 = E_2$ obviously follows that N is self-associate in the sense of Definition 5.2. Thus, N is identical with $S_{M,N}$ by Theorem 5.2.

6. In this section we present some remarks about the work of the preceding sections.

A. As a consequence of Lemma 3.4, $\overline{E}_1 \otimes \overline{E}_2 \subset \overline{E_1 \otimes E_2}$ has been proved. It does not appear to be a simple matter to describe the exact conditions imposed upon a crossnorm, under which the relation $\overline{E}_1 \otimes \overline{E}_2 = \overline{E_1 \otimes E_2}$ holds.

B. We have proved the existence of the least crossnorm, whose associate is also a crossnorm. We did not settle, however, whether the associate with every crossnorm is also a crossnorm, or there exist crossnorms whose associates are not crossnorms([7]).

C. The general crossnorm N, constructed in Theorem 5.1 by means of the greatest crossnorm N_1, we are justified to term self-associate (extending hereby Definition 5.2). Theorem 5.1.1 states that the same construction applied to any general crossnorm will always lead to a self-associate crossnorm. We did not settle, however, the problem of "uniqueness," that is, whether the same construction applied to two different general crossnorms will always lead to the same self-associate crossnorm.

In connection with the question of uniqueness let us mention the following problem.

Let N^0 and N^{00} denote two general crossnorms, and $N^0 \leqq N^{00}$. Under what conditions is $N^0 + \overline{N}^0 \geqq N^{00} + \overline{N}^{00}$?

7. The following crossnorms are worth mentioning.

DEFINITION 7.1. *Let $\sum_{i=1}^{m} f_i^0 \otimes \varphi_i^0$ be in $\mathfrak{A}(E_1, E_2)$. We put* $N(\sum_{i=1}^{m} f_i^0 \otimes \varphi_i^0) = \inf (\max_{\epsilon_i = \pm 1} \|\sum_{i=1}^{n} \epsilon_i \|f_i\| \varphi_i\|)$ *where* inf *(that is, the greatest lower bound) is taken over the set of all expressions $\sum_{i=1}^{n} f_i \otimes \varphi_i$ equivalent to $\sum_{i=1}^{m} f_i^0 \otimes \varphi_i^0$.*

THEOREM 7.1. *N in Definition 7.1 is a crossnorm.*

([7]) Added in proof: The author has since shown that if E_1 and E_2 are reflexive, that is, $\overline{\overline{E}}_i = E_i$, then the associate of every crossnorm is also a crossnorm.

Proof. It is obvious that N satisfies II and IV of Definition 3.1. We shall prove III. For two given expressions $\sum_{i=1}^{m} f_i \otimes \varphi_i$, $\sum_{i=1}^{n} g_i \otimes \psi_i$ and a given $\epsilon > 0$, we can find two expressions $\sum_{i=1}^{k} f_i' \otimes \varphi_i'$, $\sum_{i=1}^{r} g_i' \otimes \psi_i'$ such that

$$\sum_{i=1}^{k} f_i' \otimes \varphi_i' \simeq \sum_{i=1}^{m} f_i \otimes \varphi_i; \quad \max_{\epsilon_i=\pm 1} \left\| \sum_{i=1}^{k} \epsilon_i \|f_i'\| \varphi_i' \right\| \leqq N\left(\sum_{i=1}^{m} f_i \otimes \varphi_i \right) + \epsilon/2,$$

$$\sum_{i=1}^{r} g_i' \otimes \psi_i' \simeq \sum_{i=1}^{n} g_i \otimes \psi_i; \quad \max_{\epsilon_i=\pm 1} \left\| \sum_{i=1}^{r} \epsilon_i \|g_i'\| \psi_i' \right\| \leqq N\left(\sum_{i=1}^{n} g_i \otimes \psi_i \right) + \epsilon/2.$$

Let $\epsilon_i' = \pm 1$, $i=1, \cdots, k$; $\eta_j' = \pm 1$, $j=1, \cdots, r$, be chosen so that

$$\left\| \sum_{i=1}^{k} \epsilon_i' \|f_i'\| \varphi_i' + \sum_{j=1}^{r} \eta_j' \|g_j'\| \psi_j' \right\| = \max_{\epsilon_i=\pm 1;\, \eta_j=\pm 1} \left\| \sum_{i=1}^{k} \epsilon_i \|f_i'\| \varphi_i' + \sum_{i=1}^{r} \eta_i \|g_i'\| \psi_i' \right\|,$$

then IV and Definition 7.1 give

$$\begin{aligned} N\left(\sum_{i=1}^{m} f_i \otimes \varphi_i + \sum_{i=1}^{n} g_i \otimes \psi_i \right) &\leqq \left\| \sum_{i=1}^{k} \epsilon_i' \|f_i'\| \varphi_i' + \sum_{j=1}^{r} \eta_j' \|g_j'\| \psi_j' \right\| \\ &\leqq \max_{\epsilon_i=\pm 1} \left\| \sum_{i=1}^{k} \epsilon_i \|f_i'\| \varphi_i' \right\| + \max_{\epsilon_i=\pm 1} \left\| \sum_{i=1}^{r} \epsilon_i \|g_i'\| \psi_i' \right\| \\ &\leqq N\left(\sum_{i=1}^{m} f_i \otimes \varphi_i \right) + N\left(\sum_{i=1}^{n} g_i \otimes \psi_i \right) + \epsilon. \end{aligned}$$

This proves III. We shall prove VI

$$\left\| \sum_{i=1}^{n} F(f_i)\varphi_i \right\| \leqq \max_{\eta_i=\pm 1} \left\| \sum_{i=1}^{n} \eta_i |F(f_i)| \varphi_i \right\|.$$

Let $\eta_1', \cdots, \eta_n'$ denote that system of numbers 1, -1, for which the right side of the last inequality is a maximum. Then,

$$\left\| \sum_{i=1}^{n} F(f_i)\varphi_i \right\| \leqq \left\| \sum_{i=1}^{n} \eta_i' \|F\| \, \|f_i\| \varphi_i \right\| \leqq \|F\| \cdot \max_{\epsilon_i=\pm 1} \left\| \sum_{i=1}^{n} \epsilon_i \|f_i\| \varphi_i \right\|$$

and the proof of VI is a simple consequence of Lemma 2.1. Lemma 3.8 implies, therefore, I for N.

We complete the proof of the theorem by showing that N has the cross-property. Let $f \neq 0$, $\varphi \neq 0$ and $\sum_{i=1}^{n} f_i \otimes \varphi_i \simeq f \otimes \varphi$. Choose an $F \subset \overline{E}_1$ such that $F(f) = \|f\|$, $\|F\| = 1$. Lemma 2.1 gives

$$\|f\| \, \|\varphi\| = \|F(f)\varphi\| = \left\| \sum_{i=1}^{n} F(f_i)\varphi_i \right\| \leqq \|F\| \max_{\epsilon_i=\pm 1} \left\| \sum_{i=1}^{n} \epsilon_i \|f_i\| \varphi_i \right\|$$

or

$$\|f\| \, \|\varphi\| \leqq \max_{\epsilon_i=\pm 1} \left\| \sum_{i=1}^{n} \epsilon_i \|f_i\| \varphi_i \right\|.$$

This proves

$$N(f \otimes \varphi) = \|f\| \, \|\varphi\|.$$

DEFINITION 7.2. *We define a crossnorm N in $\mathfrak{A}(E_1, E_2)$ by means of the least crossnorm N_0 in $\mathfrak{A}(\overline{E}_1, \overline{E}_2)$. Let k denote a natural number. For $\sum_{i=1}^n f_i \otimes \varphi_i$ in $\mathfrak{A}(E_1, E_2)$ we put*

$$N_{(k)}\left(\sum_{i=1}^n f_i \otimes \varphi_i\right) = \sup\left(\left|\left(\sum_{j=1}^k F_j \otimes \phi_j\right)\left(\sum_{i=1}^n f_i \otimes \varphi_i\right)\right| \Big/ N_0\left(\sum_{j=1}^k F_j \otimes \phi_j\right)\right)$$

where sup (*the least upper bound*) *is taken for all sequences of k terms $F_1, \cdots, F_k$ in $\overline{E}_1$; $\phi_1, \cdots, \phi_k$ in $\overline{E}_2$.*

THEOREM 7.2. *For every natural k, $N_{(k)}$ is a crossnorm.*

Proof. I. If $\sum_{i=1}^n f_i \otimes \varphi_i \simeq 0 \otimes 0$, Lemma 2.1 gives $N_{(k)}(\sum_{i=1}^n f_i \otimes \varphi_i) = 0$. If $N_{(k)}(\sum_{i=1}^n f_i \otimes \varphi_i) = 0$, then taking $F_1 = \cdots = F_k = F$; $\phi_1 = \cdots = \phi_k = \phi$, we have $\sum_{i=1}^n f_i \otimes \varphi_i \simeq 0 \otimes 0$ by Definition 4.2 and Lemma 4.4.

That II, III, and IV hold is obvious. We shall prove that $N_{(k)}$ has the cross-property. We have

$$\begin{aligned} N_{(k)}(f^0 \otimes \varphi^0) &= \sup \left|\left(\sum_{j=1}^k F_j \otimes \phi_j\right)(f^0 \otimes \varphi^0)\right| \Big/ N_0\left(\sum_{j=1}^k F_j \otimes \phi_j\right) \\ &\leqq \sup \left\{\left|\left(\sum_{j=1}^k F_j \otimes \phi_j\right)(f^0 \otimes \varphi^0)\right| \left(\|f^0\| \, \|\varphi^0\| \Big/ \left|\left(\sum_{j=1}^k F_j \otimes \phi_j\right)(f^0 \otimes \varphi^0)\right|\right)\right\} \\ &= \|f^0\| \, \|\varphi^0\|. \end{aligned}$$

But on the other hand, putting $F_1 = \cdots = F_k = F$; $\phi_1 = \cdots = \phi_k = \phi$ we obviously get $N_{(k)}(f^0 \otimes \varphi^0) \geqq \|f^0\| \, \|\phi^0\|$. This completes the proof.

REFERENCES

1. S. Banach, *Théorie des opérations linéaires*, Warsaw, 1932.

2. D. G. Bourgin, *Closure of products of functions*, Bull. Amer. Math. Soc. vol. 46 (1940) pp. 807–815.

3. J. A. Clarkson, *Uniformly convex spaces*, Trans. Amer. Math. Soc. vol. 40 (1936) pp. 396–414.

4. M. M. Day, *Reflexive Banach spaces not isomorphic to uniformly convex spaces*, Bull. Amer. Math. Soc. vol. 47 (1941) pp. 313–317.

5. F. L. Hitchcock, *The expression of a tensor or polyadic as a sum of products*, Journal of Mathematics and Physics vol. 6 (1937) pp. 164–189.

6. ———, *Multiple invariants and generalized rank of a p-way matrix or tensor*, ibid. vol. 7 (1927) pp. 40–79.

7. D. Milman, *On some criteria for the regularity of spaces of type* (B). C. R. (Doklady) Acad. Sci. URSS. vol. 20 (1938) pp. 243–246.

8. F. J. Murray, *On complementary manifolds and projections in spaces L_p and l_p*, Trans. Amer. Math. Soc. vol. 41 (1937) pp. 138–152.

9. ———, *Bilinear transformations in Hilbert spaces*, ibid. vol. 45 (1939) pp. 474–507.

10. F. J. Murray and J. von Neumann, *On rings of operators*, Ann. of Math. (2) vol. 37 (1936) pp. 116–229.

11. J. von Neumann, *On infinite direct products*, Compositio Math. vol 6 (1938) pp. 1–77.

12. ———, *On certain topology for rings of operators*, Ann. of Math. (2) vol. 37 (1936) pp. 111–115.

13. R. Oldenburger *Composition and rank of n-way matrices and multilinear forms*, ibid vol. 35 (1934) pp. 622–657.

14. ———, *Nonsingular multilinear forms and certain p-way matrix factorizations*, Trans. Amer. Math. Soc. vol. 39 (1936) pp. 422–455.

15. B. J. Pettis, *Proof that every uniformly convex space is reflexive*, Duke Math. J. vol. 5 (1939) pp. 249–253.

16. H. Weyl, *The theory of groups and quantum-mechanics*, translated from German by H. P. Robertson, New York, 1931.

COLUMBIA UNIVERSITY,
NEW YORK, N. Y.

DIRECT METHODS IN GEOMETRICAL OPTICS

BY

M. HERZBERGER

Communication no. 851 from the Kodak Research Laboratories

This paper presents a different approach to the problems of geometrical optics, in order to attack some problems hitherto insoluble in practice.

The investigation has been restricted to rotationally symmetric optical systems for practical reasons only; the procedures can be applied to the general case. Let us choose two rectangular coordinate systems such that the z- and z'-axes coincide with the optical axis, the y- and y'-, and x- and x'-axes being, respectively, parallel. An object ray is then defined by its intersection point with the plane $z=0$; (vector $a^{\rightarrow}$, coordinates $x, y, 0$) and by the vector $s^{\rightarrow}$ of length n (n refractive index) along the ray (coordinates $\xi, \eta, \zeta=(n^2-\xi^2-\eta^2)^{1/2}$); the image is given by $a'^{\rightarrow}(x', y', 0)$ and $s'^{\rightarrow}(\xi', \eta', \zeta'=(n'^2-\xi'^2-\eta'^2)^{1/2})$, with n' the refractive index of the image space.

The problem of the optical designer is to find the image ray if the object ray is given, or in other words, to compute, for a given optical system, x', y', ξ', and η' as functions of x, y, ξ, and η. If this problem is solved for a single surface and arbitrary positions of the object and image planes, it is solved for any rotationally symmetrical optical system merely by making a succession of substitutions. To solve the problem for a single surface, which will here be assumed to be a spherical surface, we first place object and image planes at the center of the refracting surface, and then calculate the functions. Having done this, we have only to compute the intersection points of the image rays with a parallel plane through another origin, a simple geometric problem. We can use this method for the manifold of all rays in a procedure similar to the ordinary way of tracing meridional rays; and applied to an individual ray, it becomes a straightforward method for tracing skew rays through an optical system.

I. General formulae

Before deriving these equations, we shall inspect the general conditions of *optical image formation*(1). Because of the rotational symmetry we can write

---

Presented to the Society, October 31, 1942; received by the editors April 17, 1942.

(1) The term "optical image formation" refers to the one-to-one correspondence of object and image rays in an optical system. Not every one-to-one correspondence can be considered as an optical image formation. The conditions are derived in this paper: For the mathematician it might be noted that an optical image formation is a special kind of contact transformation.

$$
(1) \qquad \begin{aligned} x' &= Ax + B\xi, & \xi' &= Cx + D\xi \\ y' &= Ay + B\eta, & \eta' &= Cy + D\eta \end{aligned}
$$

in which A, B, C, and D are functions of u_1, u_2, and u_3, which are symmetric functions of x, y, ξ, and η:

$$
(2) \qquad \begin{aligned} u_1 &= (1/2)(x^2 + y^2), \\ u_2 &= x\xi + y\eta, \\ u_3 &= (1/2)(\xi^2 + \eta^2). \end{aligned}
$$

However, A, B, C, and D are not arbitrary functions of u_1, u_2, and u_3. Our first task is to derive the differential equations connecting them.

According to the fundamental optical invariant(²), for any two-parameter manifolds (parameters u, v), we have (abbreviating $\partial \vec{s}/\partial u = \vec{s}_u$),

$$
(3) \qquad \vec{s}\,'_u \vec{a}\,'_v - \vec{s}\,'_v \vec{a}\,'_u = \vec{s}_u \vec{a}_v - \vec{s}_v \vec{a}_u .
$$

Taking the following variables in turn for u and v: x, y; x, ξ; x, η; y, ξ; y, η; and ξ, η and turning from the vectors to the coordinates, we find that

$$
(4) \qquad \begin{aligned} &(x'_x \xi'_y + y'_x \eta'_y) - (x'_y \xi'_x + y'_y \eta'_x) = 0, \\ &(x'_x \xi'_\xi + y'_x \eta'_\xi) - (x'_\xi \xi'_x + y'_\xi \eta'_x) = 1, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \end{aligned}
$$

Writing A_ν for $\partial A/\partial u_\nu$, and so forth, we now obtain from (1) and (2),

$$
(5) \qquad \begin{aligned} x'_x &= A + A_1 x^2 + (A_2 + B_1)x\xi + B_2 \xi^2, \\ x'_y &= A_1 xy + A_2 x\eta + B_1 \xi y + B_2 \xi\eta, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \end{aligned}
$$

Inserting (5) in (4), we obtain the equations

$$
(6) \qquad \begin{aligned} (x\eta - y\xi)I &= 0, \\ (AD - BC) + x^2 I + x\xi II + \xi^2 III &= 1, \\ xyI + x\eta II + \xi\eta III &= 0, \\ xyI + y\xi II + \xi\eta III &= 0, \\ (AD - BC) + y^2 I + y\eta II + \eta^2 III &= 1, \\ (x\eta - y\xi)III &= 0; \end{aligned}
$$

(²) For the historical background of formula (3) and its connection with different branches of mathematics, the reader is referred to M. Herzberger, *Theory of transversal curves and the connections between the calculus of variations and the theory of partial differential equations*, Proc. Nat. Acad. Sci. U.S.A. vol. 24 (1938) pp. 466–473.

where

$$
(7)\qquad
\begin{aligned}
I &= AC_2 + DA_1 - CA_2 - BC_1 + 2u_1(A_1C_2 - A_2C_1)\\
&\quad + u_2(A_1D_2 - A_2D_1 + B_1C_2 - B_2C_1) + 2u_3(B_1D_2 - B_2D_1),\\
II &= AC_3 + DB_1 - CA_3 - BD_1 + 2u_1(A_1C_3 - A_3C_1)\\
&\quad + u_2(A_1D_3 - A_3D_1 + B_1C_3 - B_3C_1) + 2u_3(B_1D_3 - B_3D_1),\\
III &= AD_3 + DB_2 - CB_3 - BD_2 + 2u_1(A_2C_3 - A_3C_2)\\
&\quad + u_2(A_2D_3 - A_3D_2 + B_2C_3 - B_3C_2) + 2u_3(B_2D_3 - B_3D_2).
\end{aligned}
$$

Equations (6) can be satisfied for all rays only if

$$
(8)\qquad
\begin{aligned}
I = II = III = 0;\\
AD - BC = 1.
\end{aligned}
$$

The fulfillment of the four equations (8) is necessary and sufficient to guarantee that equation (1) determines an optical image formation. We can of course eliminate one of the functions, for example, D, and have three functions A, B, and C, and three differential equations connecting them. We find that

$$
(9)\qquad
\begin{aligned}
D &= \frac{1 + BC}{A},\\
D_\nu &= \frac{C}{A} B_\nu + \frac{B}{A} C_\nu - \frac{1 + BC}{A^2} A_\nu.
\end{aligned}
$$

Inserting this in (7), we obtain these equations:

$$
(10)\qquad
\begin{aligned}
&(AC_2 - CA_2) + \frac{B}{A}(CA_1 - AC_1) + \frac{A_1}{A} + (A_1C_2 - A_2C_1)\left(2u_1 + \frac{B}{A}u_2\right)\\
&\qquad + (A_1B_2 - A_2B_1)\left(\frac{C}{A}u_2 + 2u_3\frac{1 + BC}{A^2}\right)\\
&\qquad + (B_1C_2 - B_2C_1)\left(u_2 + 2u_3\frac{B}{A}\right) = 0,\\
&(AC_3 - CA_3) + \frac{B^2}{A}(CA_1 - AC_1) + \frac{BA_1 + AB_1}{A^2}\\
&\qquad + (A_1C_3 - A_3C_1)\left(2u_1 + \frac{B}{A}u_2\right)\\
&\qquad + (A_1B_3 - A_3B_1)\left(\frac{C}{A}u_2 + 2u_3\frac{1 + BC}{A^2}\right)\\
&\qquad + (B_1C_3 - B_3C_1)\left(u_2 + 2u_3\frac{B}{A}\right) = 0,
\end{aligned}
$$

$$
(10)\qquad
\begin{aligned}
&\frac{B}{A}(AC_3 - CA_3) + \frac{B^2}{A^2}(CA_2 - AC_2) + \frac{BA_2 + AB_2}{A^2} - \frac{A_3}{A} \\
&\qquad + (A_2C_3 - A_3C_2)\left(2u_1 + \frac{B}{A}u_2\right) \\
&\qquad + (A_2B_3 - A_3B_2)\left(\frac{C}{A}u_2 + 2u_3\frac{1+BC}{A^2}\right) \\
&\qquad + (B_2C_3 - B_3C_2)\left(u_2 + 2u_3\frac{B}{A}\right) = 0.
\end{aligned}
$$

Let us multiply these equations by $A_3, -A_2, A_1; B_3, -B_2, B_1; C_3, -C_2, C_1$, respectively, and add. We thus obtain three new equations, which can replace equations (10) if

$$
(11)\qquad
\begin{vmatrix} A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \\ C_1 & C_2 & C_3 \end{vmatrix} = \Delta \neq 0;
$$

that is, A, B, and C are three independent functions. (In this case we can construct their inverse functions, that is, we can calculate u_1, u_2, and u_3 as functions of A, B, and C. We shall make use of this later.) In either case, we obtain from (10) three equations:

$$
(12)\qquad
\begin{aligned}
&A(C_2A_3 - C_3A_2) + B(A_1C_3 - C_1A_3) + \frac{B^2}{A}(C_1A_2 - C_2A_1) \\
&\qquad + \frac{1}{A}(B_2A_1 - B_1A_2) + \left(u_2 + 2u_3\frac{B}{A}\right)\Delta = 0, \\
&A(C_2B_3 - C_3B_2) + B(B_1C_3 - C_1B_3) + C(A_3B_2 - A_2B_3) \\
&\qquad + \frac{1+BC}{A}(A_1B_3 - B_1A_3) \\
&\qquad + \frac{B}{A}\left(\frac{1+BC}{A}\right)(A_2B_1 - A_1B_2) + \frac{B^2}{A}(C_1B_2 - C_2B_1) \\
&\qquad + \left(2u_1 + \frac{B}{A}u_2\right)\Delta = 0, \\
&C(C_2A_3 - C_3A_2) + \frac{1+BC}{A}(A_1C_3 - C_1A_3) \\
&\qquad + \frac{B}{A}(A_2C_1 - A_1C_2)\left(\frac{1+BC}{A}\right) + \frac{1}{A}(B_2C_1 - B_1C_2) \\
&\qquad + \left(\frac{C}{A}u_2 + 2u_3\frac{1+BC}{A^2}\right)\Delta = 0.
\end{aligned}
$$

These equations are equivalent to:

$$
(13)\quad
\begin{aligned}
&A(A_1C_3 - A_3C_1) + B(A_2C_1 - A_1C_2) + C(A_2B_1 - A_1B_2)\\
&\qquad\qquad\qquad + A(B_2C_1 - B_1C_2) + 2u_3\Delta = 0,\\
&A(A_3C_2 - A_2C_3) + \frac{1 + BC}{A}(A_2B_1 - A_1B_2) - B(B_2C_1 - B_1C_2) + u_2\Delta = 0,\\
&A(B_3C_2 - C_3B_2) + B(B_1C_3 - C_1B_3) + C(A_3B_2 - A_2B_3)\\
&\qquad - B(A_3C_2 - C_3A_2) + \frac{1 + BC}{A}(A_1B_3 - B_1A_3)\\
&\qquad + 2u_1\Delta = 0.
\end{aligned}
$$

If Δ is not identically equal to zero, we now get very simple equations for the inverse functions, if we insert

$$
(14)\quad
\begin{aligned}
&\frac{\partial u_1}{\partial A} = \frac{B_2C_3 - B_3C_2}{\Delta}, &&\frac{\partial u_1}{\partial B} = \frac{A_3C_2 - A_2C_3}{\Delta}, &&\frac{\partial u_1}{\partial C} = \frac{A_2B_3 - A_3B_2}{\Delta},\\
&\frac{\partial u_2}{\partial A} = \frac{B_1C_3 - B_3C_1}{\Delta}, &&\frac{\partial u_2}{\partial B} = \frac{A_1C_3 - A_3C_1}{\Delta}, &&\frac{\partial u_2}{\partial C} = \frac{A_3B_1 - A_1B_3}{\Delta},\\
&\frac{\partial u_3}{\partial A} = \frac{B_1C_2 - B_2C_1}{\Delta}, &&\frac{\partial u_3}{\partial B} = \frac{A_2C_1 - A_1C_2}{\Delta}, &&\frac{\partial u_3}{\partial C} = \frac{A_1B_2 - A_2B_1}{\Delta}.
\end{aligned}
$$

Thus, equations (13) give finally

$$
(15)\quad
\begin{aligned}
A\frac{\partial u_2}{\partial B} + B\frac{\partial u_3}{\partial B} - C\frac{\partial u_3}{\partial C} - A\frac{\partial u_3}{\partial A} + 2u_3 &= 0,\\
A\frac{\partial u_1}{\partial B} - \frac{1 + BC}{A}\cdot\frac{\partial u_3}{\partial C} + B\frac{\partial u_3}{\partial A} + u_2 &= 0,\\
A\frac{\partial u_1}{\partial A} - B\frac{\partial u_2}{\partial A} + C\frac{\partial u_1}{\partial C} + B\frac{\partial u_1}{\partial B} + \frac{1 + BC}{A}\cdot\frac{\partial u_2}{\partial C} - 2u_1 &= 0,
\end{aligned}
$$

as fundamental equations for optical image formation.

II. A special kind of optical image formation

As an example, let us consider a special case in which $\Delta \equiv 0$. Let us assume $B \equiv 0$; then

$$
(16)\quad
\begin{aligned}
x' &= Ax, & \xi' &= Cx + (1/A)\xi,\\
y' &= Ay, & \eta' &= Cy + (1/A)\eta.
\end{aligned}
$$

This kind of image formation is of great importance in optics. A single

sphere produces this type of image; so does a system of concentric surfaces, a so-called concentric lens system. On the other hand, the attempt to realize the dream of an optical designer, a system which gives a sharp image of one plane $z=0$ upon another ($z'=0$), leads to an image formation given by formula (16) with the special condition that A is a constant.

In general, A must be different from 0 and, inserting $B=B_1=B_2=B_3=0$, equations (10) give

$$\begin{aligned} (AC_2 - CA_2) + (A_1/A) + (A_1C_2 - A_2C_1)2u_1 &= 0, \\ (AC_3 - CA_3) \qquad\qquad\qquad + (A_1C_3 - A_3C_1)2u_1 &= 0, \\ - (A_3/A) + (A_2C_3 - A_3C_2)2u_1 &= 0, \end{aligned} \tag{17}$$

as differential equations for the two functions. Equations (12) give

$$A(C_2A_3 - C_3A_2) = 0 \quad \text{and} \quad C(C_2A_3 - C_3A_2) + \frac{1}{A}(A_1C_3 - C_1A_3) = 0, \tag{18}$$

or

$$C_2A_3 - C_3A_2 = A_1C_3 - A_3C_1 = 0. \tag{19}$$

Inserting (19) in (17), we find that

$$\begin{gathered} A_3 = C_3 = 0, \\ (AC_2 - CA_2) + (A_1/A) + (A_1C_2 - A_2C_1)2u_1 = 0. \end{gathered} \tag{20}$$

We see that in this case A and C are functions of u_1 and u_2 alone. We introduce $D=1/A$ and find, as differential equations for C and D,

$$\begin{gathered} D_3 = C_3 = 0, \\ DC_2 + CD_2 - DD_1 + (C_1D_2 - D_1C_2)2u_1 = 0. \end{gathered} \tag{21}$$

In the special case of $D=\text{const.}$, $C_2=0$. That means C is a function of u_1 alone. We have

$$\begin{aligned} x' &= A_0x, & \xi' &= C(u_1)x + (1/A_0)\xi, \\ y' &= A_0y, & \eta' &= C(u_1)y + (1/A_0)\eta. \end{aligned} \tag{22}$$

Equation (22) is a generalization of the well known sine condition of Abbe; A_0 is the magnification of the image and $(22)_2$ can be written in the familiar form

$$A_0\xi' - \xi = C(u_1)x, \qquad A_0\eta' - \eta = C(u_1)y, \tag{23}$$

where the right sides of equations (23) are independent of ξ and η.

Let us now investigate the general case, in which C and D are independent functions of u_1 and u_2, such that

$$C_1D_2 - C_2D_1 \neq 0. \tag{24}$$

Here we may again construct the inverse functions, and consider u_1 and u_2 as functions of C and D. Then we have

$$\begin{aligned} \frac{\partial u_1}{\partial C} &= \frac{D_2}{C_1D_2 - C_2D_1}, & \frac{\partial u_1}{\partial D} &= \frac{-C_2}{C_1D_2 - C_2D_1}, \\ \frac{\partial u_2}{\partial C} &= \frac{-D_1}{C_1D_2 - C_2D_1}, & \frac{\partial u_2}{\partial D} &= \frac{C_1}{C_1D_2 - C_2D_1}. \end{aligned} \tag{25}$$

Inserting (25) in (21) gives

$$C\frac{\partial u_1}{\partial C} - D\frac{\partial u_1}{\partial D} + D\frac{\partial u_2}{\partial C} + 2u_1 = 0. \tag{26}$$

The solution of this differential equation would solve our problem.

III. The single spherical surface

Again we attack a special case of the aforesaid problem, namely, the single refracting spherical surface. The coordinate origins in the object and image spaces are placed at the center of the sphere, the x- and x'-, y- and y'-axes coinciding. Let $\vec{a}$ ($\vec{a'}$) be the vector $(x, y, 0)$ $(x', y', 0)$, respectively, to the intersection point of the ray with the coordinate plane $z(z')=0$; let $\vec{s}(\xi, \eta, \zeta)$, $\vec{s'}(\xi', \eta', \zeta')$, respectively, be the vectors of length n (n') along the object and image rays.

Let r be the radius of the refracting surface, positive if the surface is convex with respect to the direction of the light ray, and negative if the surface is concave. Let $\vec{r}$ be the vector of length r along the incidence normal. The refraction law can then be written

$$\vec{s} \times \vec{r} = \vec{s'} \times \vec{r}, \quad \text{or} \quad \vec{s'} - \vec{s} = C\vec{r}. \tag{27}$$

The coordinates are then given by equations (2):

$$\begin{aligned} u_1 &= (x^2 + y^2)/2 = \vec{a}^2/2, \\ u_2 &= x\xi + y\eta = \vec{a}\cdot\vec{s}. \end{aligned} \tag{28}$$

Two values, λ and λ', exist such that

$$\vec{a} + \lambda\vec{s} = \vec{r} = \vec{a'} + \lambda'\vec{s'}, \tag{29}$$

and $\vec{a'}$ and $\vec{a}$ must have the same direction, since they both lie in the intersection line of the incidence plane with the plane $z=0$. From (29) and (27) we find that

$$\vec{a'} \times \vec{s'} = \vec{r} \times \vec{s'} = \vec{r} \times \vec{s} = \vec{a} \times \vec{s}. \tag{30}$$

This gives finally

$$(31)\qquad \begin{aligned} a'^{\rightarrow} &= (1/D)a^{\rightarrow}, \\ s'^{\rightarrow} &= Ca^{\rightarrow} + Ds^{\rightarrow}, \end{aligned}$$

where C is given by equation $(27)_2$; $a^{\rightarrow}\times s^{\rightarrow}$ is an invariant vector for refraction. Its direction is perpendicular to the incident plane; its length p is the length of the perpendicular dropped from the center to the incident (refracted) ray, multiplied by the corresponding refractive index.

We have

$$(32)\qquad p^2 = (a^{\rightarrow} \times s^{\rightarrow})^2 = a^{\rightarrow 2}s^{\rightarrow 2} - (a^{\rightarrow}\cdot s^{\rightarrow})^2 = 2n^2u_1 - u_2^2.$$

Equation (31) now gives

$$(33)\qquad \begin{aligned} s'^{\rightarrow} \times s^{\rightarrow} &= C(a^{\rightarrow} \times s^{\rightarrow}), \\ s'^{\rightarrow}\cdot s^{\rightarrow} &= Ca^{\rightarrow}\cdot s^{\rightarrow} + Ds^{\rightarrow 2}. \end{aligned}$$

If δ is the angle between s and s' (the angle of deviation), equations (33) are equivalent to

$$nn' \sin \delta = Cp, \qquad nn' \cos \delta = Cu_2 + Dn^2,$$

or

$$(34)\qquad C = \frac{nn' \sin \delta}{p}, \qquad D = \frac{1}{n^2}\left\{nn' \cos \delta - \frac{nn' \sin \delta}{p} u_2\right\}.$$

Our remaining problem is to express δ as a function of p, and then, by using (32), as a function of u_1 and u_2. Equation $(27)_2$ gives

$$(35)\qquad Cr^{\rightarrow 2} = s'^{\rightarrow}\cdot r^{\rightarrow} - s^{\rightarrow}\cdot r^{\rightarrow},$$

or, because of (30),

$$(36)\qquad C = (1/r)\{(n'^2 - (p/r)^2)^{1/2} - (n^2 - (p/r)^2)^{1/2}\},$$

$$(37)\qquad \begin{aligned} \sin \delta &= (p/nn'r)\{(n'^2 - (p/r)^2)^{1/2} - (n^2 - (p/r)^2)^{1/2}\}, \\ \cos \delta &= (p^2/nn'r^2) + \frac{1}{nn'}([n'^2 - (p/r)^2][n^2 - (p/r)^2])^{1/2}. \end{aligned}$$

Inserting (37) in (34) gives C and D as functions of p and u_2, and therefore, because of (32), as functions of u_1 and u_2.

To solve the reverse problem, that is, to calculate u_1 and u_2 as functions of C and D, we proceed as follows:

Equation (37) gives

$$(38a)\qquad (\cos \delta - p^2/nn'r^2)^2 = (1/nn')^2(n'^2 - p^2/r^2)(n^2 - p^2/r^2),$$

$$(nn' \sin \delta)^2 + (p^2/r^2)(2nn' \cos \delta - n^2 + n'^2) = 0; \tag{38b}$$

or, considering $(34)_1$

$$2nn' \cos \delta = n^2 + n'^2 - C^2r^2. \tag{39}$$

Inserting this in $(34)_2$, we obtain

$$u_2 = \frac{n^2 + n'^2 - C^2r^2}{2C} - \frac{D}{C} n^2. \tag{40}$$

Equation (34) gives

$$\begin{aligned} p^2 = \left(\frac{nn' \sin \delta}{C}\right)^2 &= \frac{n^2n'^2}{C^2} - \left(\frac{n^2 + n'^2 - C^2r^2}{2C}\right)^2 = 2n^2u_1 - u_2^2 \\ &= 2n^2u_1 - \left(\frac{n^2 + n'^2 - C^2r^2}{2C} - \frac{D}{C} n^2\right)^2, \end{aligned} \tag{41}$$

which gives

$$\begin{aligned} u_1 &= - Dr^2/2 + \frac{1 - D}{2C^2}(n'^2 - n^2D), \\ u_2 &= - Cr^2/2 + \frac{n^2(1 - D) + (n'^2 - n^2D)}{2C}. \end{aligned} \tag{42}$$

We see that u_1 and u_2 are rational functions of C and D, fulfilling equation (26).

Equations (42) are very valuable for calculating the coefficients of C and D, written as a series in u_1 and u_2.

IV. The plane surface

The equations for refracting a ray at a plane surface are very simple. We let the z- and z'-axes coincide with the axis of the system, and place the origins at the intersection point of the axis with the plane surface. Applying the refraction law here gives

$$\begin{aligned} x' &= x, \qquad \xi' = \xi, \\ y' &= y, \qquad \eta' = \eta. \end{aligned} \tag{43}$$

This means that the transformation is the identical transformation.

V. Transition formulae

To obtain the formulae for tracing a system of rays through a system of centered lenses, we must also know how to find, from the coordinates of the intersection point of the image ray with the reference plane through the

center of one surface, the coordinates of the ray intersection in the reference plane through the center of the succeeding surface. We call the distance between the two centers m. From analytical geometry, we get

$$(44)\qquad \begin{aligned} x' &= x + \frac{m}{(1 - 2u_3/n^2)^{1/2}}\xi, & \xi' &= \xi, \\ y' &= y + \frac{m}{(1 - 2u_3/n^2)^{1/2}}\eta, & \eta' &= \eta. \end{aligned}$$

In tracing rays, we can use the formulae described in the preceding section. We can simplify the calculation since, instead of tracing x, x', y' and y, it proves to be sufficient to trace u_1, u_2, and u_3, because the functions depend only upon these variables. Moreover, for any optical image formation, we have the relation

$$(45)\qquad u_1'u_3' - u_2'^2 = (x'\eta' - y'\xi')^2 = (x\eta - y\xi)^2 = u_1u_3 - u_2^2,$$

which follows directly from (1). The actual ray tracing formulae will be published later.

VI. Image error theory

Finally, let us sketch briefly how the image error theory can be derived, using this new method of approach.

Let us develop A, B, C, and D as functions of u_1, u_2, and u_3, into a series and inspect the equations for small values of u_1, u_2, and u_3.

If we assume u_1, u_2, and u_3 to be negligible, we obtain Gaussian optics, within the realm of which A, B, C, and D can be regarded as constant values. We have

$$(46)\qquad \begin{aligned} x' &= A_0x + B_0\xi, & \xi' &= C_0x + D_0\xi, \\ y' &= A_0y + B_0\eta, & \eta' &= C_0y + D_0\eta, \end{aligned}$$

with $(8)_2$

$$A_0D_0 - B_0C_0 = 1.$$

The special case, that object and image are in optically conjugated points, is indicated by

$$(47)\qquad B_0 = 0.$$

Then $D_0 = 1/A_0$, and

$$(48)\qquad \begin{aligned} x' &= A_0x, & \xi' &= C_0x + (1/A_0)\xi, \\ y' &= A_0y, & \eta' &= C_0y + (1/A_0)\eta. \end{aligned}$$

From (48) and (46) can be derived all the laws of Gaussian optics.

Let us now consider the linear members of A, B, C, and D, but neglect all higher orders. This leads to the so-called Seidel theory of aberrations. We have again

$$D = \frac{1 + BC}{A}, \tag{49}$$

but the nine first-order derivatives of A, B, and C are not independent. Equations (9) show that between these derivatives at $u_1 = u_2 = u_3 = 0$ there exist the three equations

$$\begin{aligned} &\frac{B_0}{A_0}(A_0C_1 - C_0A_1) - (A_0C_2 - C_0A_2) = \frac{A_1}{A_0}, \\ &\frac{B_0^2}{A_0}(A_0C_1 - C_0A_1) - (A_0C_3 - C_0A_3) = \frac{B_0A_1 + B_1A_0}{A_0^2}, \\ &\frac{B_0^2}{A_0^2}(A_0C_2 - C_0A_2) - \frac{B_0}{A_0}(A_0C_3 - C_0A_3) = \frac{B_0A_2 - B_2A_0}{A_0^2} - \frac{A_3}{A_0}. \end{aligned} \tag{50}$$

From these equations, we get

$$\begin{aligned} B_0(B_1 - A_2) &= A_0(B_2 - A_3), \\ A_0C_2 - C_0A_2 &= \frac{B_0}{A_0}(A_0C_1 - C_0A_1) - \frac{A_1}{A_0}, \\ A_0C_3 - C_0A_3 &= \frac{B_0^2}{A_0}(A_0C_1 - C_0A_1) - \frac{B_0A_1 - B_1A_0}{A_0^2}. \end{aligned} \tag{51}$$

Let us now put the origin at the Gaussian image point, so that $B_0 = 0$, and equations (51) simplify to

$$\begin{aligned} B_2 &= A_3, \\ A_0C_2 - C_0A_2 &= -\frac{A_1}{A_0}, \\ A_0C_3 - C_0A_3 &= -\frac{B_1}{A_0}. \end{aligned} \tag{52}$$

Substituting this in the first equation of (1), we find, after rearranging, that

$$\begin{aligned} x' - A_0x &= (A_1u_1 + A_2u_2 + A_3u_3)x + (B_1u_1 + B_2u_2 + B_3u_3)\xi, \\ y' - A_0y &= (A_1u_1 + A_2u_2 + A_3u_3)y + (B_1u_1 + B_2u_2 + B_3u_3)\eta. \end{aligned} \tag{53}$$

Geometrical inspection will show that the five quantities A_1, A_2, $A_3=B_2$, B_1, and B_3 correspond to the five image errors:

B_3 represents the spherical aberration;

$A_3=B_2$ represents the coma error;

A_2 and B_1 represent the field errors;

A_1 represents the distortion.

In (53) the coordinates x', y' of the intersections of the image ray with the image plane are given as functions of x, y, ξ, η. $x=y=0$ characterizes the axis point of the object, $\xi=\eta=0$ characterizes the infinite point. We can say, therefore, that A_1, A_2, A_3, B_1, B_2, B_3 give us the image errors of our object for a stop at infinity. To obtain the image errors for a finite stop, we have to replace ξ, η, by the coordinates, x_p, y_p, of the intersection point of our ray with the plane of the diaphragm. Within the region of validity of the Seidel theory, we get a simple linear transformation. If k is the distance between object and stop, we have

$$x_p = x + k\xi, \qquad \xi = \frac{x_p - x}{k},$$

(54) or

$$y_p = y + k\eta, \qquad \eta = \frac{y_p - y}{k}.$$

Inserting (54) into (53) we find x' and y' as functions of x, y, x_p, y_p, the coefficients being the image errors for finite stop.

The method developed in this paper allows one to obtain the image coordinates in a rotation symmetric optical system as functions of the object coordinates by a series of substitutions. The only other general method having Hamilton's characteristic function leads to an elimination problem, hitherto unsolved.

Hamilton's method is more elegant since it uses only a single function to describe an optical instrument; the method of this paper leads to four functions connected by three differential equations. However, an explicit way was found to construct our function for any given system of centered lenses, whereas the characteristic function of Hamilton is explicitly known only for a single refracting surface or a plane parallel plate. Thus, the new method seems to be more adaptable to practical problems.

The last paragraph tries to show that the access to the image theory by the direct method is as simple as it is by using Hamilton's characteristic function.

Rochester, N. Y.

MEAN-VALUES AND HARMONIC POLYNOMIALS

BY
E. F. BECKENBACH AND MAXWELL READE

Introduction

0.1. It is well known that if the function $f(x, y)$ is harmonic in a finite domain (non-null connected open set) D, then at each point (x_0, y_0) in D, $f(x, y)$ satisfies the equation

$$(1) \qquad f(x_0, y_0) = \frac{1}{\pi r^2} \iint_{D(x_0, y_0; r)} f(x_0 + \xi, y_0 + \eta) d\xi d\eta$$

for each circular disc

$$D(x_0, y_0; r): \quad \xi^2 + \eta^2 = (x - x_0)^2 + (y - y_0)^2 \leqq r^2$$

lying in D. Conversely, if $f(x, y)$ is superficially summable in the interior of a finite domain D, and if (1) holds for each point (x_0, y_0) and each disc $D(x_0, y_0; r)$ about (x_0, y_0) in D, then $f(x, y)$ is harmonic in D(1).

It follows that (1) may be taken as the defining equation for harmonic functions.

0.2. Similarly, if $f(x, y)$ is superficially summable in the interior of a finite simply-connected domain D, and if $f(x, y)$ is summable on each circle

$$C(x_0, y_0; r): \quad \xi^2 + \eta^2 = (x - x_0)^2 + (y - y_0)^2 = r^2$$

lying in D, then a necessary and sufficient condition that $f(x, y)$ be harmonic in D is that at each point (x_0, y_0) in D, $f(x, y)$ satisfy the equation

$$(2) \qquad f(x_0, y_0) = \frac{1}{2\pi r} \int_{C(x_0, y_0; r)} f(x_0 + \xi, y_0 + \eta) ds$$

for each circle $C(x_0, y_0; r)$ in D.

As with (1), (2) may be taken as the defining equation for harmonic functions.

0.3. The following theorem is analogous to a result of Beckenbach and Rado concerning subharmonic functions(2).

THEOREM 1. *If $f(x, y)$ is continuous in a finite domain D, then a necessary*

Presented to the Society, December 30, 1941; received by the editors March 20, 1942.

(1) See E. Levi, *Supra una proprietà caratteristica delle funzione armoniche*, Atti della Reale Academia Lincei vol. 18 (1909) pp. 10–15, and L. Tonelli, *Sopra una proprietà caratteristica delle funzione armoniche*, ibid. pp. 577–582.

(2) E. F. Beckenbach and Tibor Rado, *Subharmonic functions and surfaces of negative curvature*, Trans. Amer. Math. Soc. vol. 35 (1933) pp. 662–674.

and sufficient condition that $f(x, y)$ be harmonic in D is that for each point (x_0, y_0) in D, the equation

$$(3)\quad \frac{1}{2\pi r}\int_{C(x_0,y_0;r)} f(x_0+\xi, y_0+\eta)ds = \frac{1}{\pi r^2}\iint_{D(x_0,y_0;r)} f(x_0+\xi, y_0+\eta)d\xi d\eta$$

hold for all $D(x_0, y_0; r)$ in D.

Proof. If $f(x, y)$ is harmonic in D, then (3) follows from (1) and (2).

To prove that (3) is a sufficient condition that $f(x, y)$ be harmonic in D, we consider the circular average([3])

$$f(x, y; \rho) \equiv \frac{1}{\pi\rho^2}\iint_{D(x,y;\rho)} f(x+\xi, y+\eta)d\xi d\eta,$$

which is defined in an open subset of D. For $D(x_0, y_0; \rho)$ in D, a computation yields

$$\frac{d}{d\rho} f(x_0, y_0; \rho) = \frac{2}{\rho}\left[\frac{1}{2\pi\rho}\int_{C(x_0,y_0;\rho)} f(x_0+\xi, y_0+\eta)ds - \frac{1}{\pi\rho^2}\iint_{D(x_0,y_0;\rho)} f(x_0+\xi, y_0+\eta)d\xi d\eta\right],$$

which, with (3), shows that $f(x_0, y_0; \rho)$ is independent of ρ. But since $f(x, y)$ is continuous, we have $f(x, y; \rho) \rightrightarrows f(x, y)$, as $\rho \to 0$, on each closed subset of D, so that

$$f(x_0, y_0) = f(x_0, y_0; r) = \frac{1}{\pi r^2}\iint_{D(x_0,y_0;r)} f(x_0+\xi, y_0+\eta)d\xi d\eta$$

for each $D(x_0, y_0; r)$ in D. Therefore $f(x, y)$ is harmonic in D.

If we should assume only that $f(x, y)$ is superficially summable and satisfies (3), then it would not follow that $f(x, y)$ is harmonic; consider, for example, the function which vanishes identically except at the origin, where it assumes the value 1.

0.4. The right-hand members in (1) and (2) are areal and peripheral averages (mean-values), respectively; in each instance the range of integration is circular. The question arises as to the nature of the functions which are defined by relations similar to (1), (2) and (3), in which the geometric figure is square, elliptic, and so on([4]).

In this paper we delineate the classes of functions defined by the condition

---

([3]) See § 1.3 below.

([4]) Cf. W. Brödel, *Funktionen mit Gaussischer Mittelwerteigenschaft für konvexe Kurven und Bereiche*, Deutsche Mathematik, vol. 4 (1939) pp. 3–15. By combining our methods with his, one can simplify the proofs of some of his results concerning general mean-values.

that their averages over regular polygons of n sides satisfy conditions similar to (1), (2) and (3)(5).

0.5. Since a circle may be considered as a limit of a sequence of circumscribed (or inscribed) regular polygons, the results of this introductory section may be considered as limiting cases of some of the results obtained below.

1. Lemmas

1.1. We recall that if $n \geqq 2$, then for all angles ψ we have

$$\sum_{m=0}^{n-1} \sin\left(\psi + \frac{2\pi m}{n}\right) = \sum_{m=0}^{n-1} \cos\left(\psi + \frac{2\pi m}{n}\right) = 0, \tag{4}$$

$$\sum_{m=0}^{n-1} \sin\left(\psi + \frac{2\pi m}{n}\right) \cos\left(\psi + \frac{2\pi m}{n}\right) = 0, \tag{5}$$

$$\sum_{m=0}^{n-1} \sin^2\left(\psi + \frac{2\pi m}{n}\right) = \sum_{m=0}^{n-1} \cos^2\left(\psi + \frac{2\pi m}{n}\right) = n/2. \tag{6}$$

Thus

$$\sum_{m=0}^{n-1}\left[\cos\left(\psi + \frac{2\pi m}{n}\right) + i \sin\left(\psi + \frac{2\pi m}{n}\right)\right]$$
$$= (\cos\psi + i\sin\psi)\sum_{m=0}^{n-1}\left(\cos\frac{2\pi m}{n} + i\sin\frac{2\pi m}{n}\right),$$

and this vanishes since the sum of the nth roots of unity is zero. The remaining formulas in (4), (5) and (6) can be established in a similar way.

More generally, for $n \geqq 1$,

$$\sum_{m=0}^{n-1}\left[\cos\left(\psi + \frac{2\pi km}{n}\right) + i \sin\left(\psi + \frac{2\pi km}{n}\right)\right] = n\delta_{k,n}(\cos\psi + i\sin\psi), \tag{7}$$

where $\delta_{k,n} = 1$ if k is an integral multiple of n, and where $\delta_{k,n} = 0$ otherwise(6).

1.2. $P_n(x_0, y_0; r; \phi)$, $n \geqq 3$, shall denote the closed finite region bounded by the regular n-gon $p_n(x_0, y_0; r; \phi)$, whose center is at (x_0, y_0) and whose inscribed circle has radius r; ϕ denotes the angle from R to N, $-\pi/n \leqq \phi < \pi/n$, where R is the ray extending horizontally to the right from (x_0, y_0) and N is the exterior normal at the point where R emerges from the polygon. $|P_n(x_0, y_0; r; \phi)|$ shall denote the area of $P_n(x_0, y_0; r; \phi)$ and $|p_n(x_0, y_0; r; \phi)|$ shall denote the length of $p_n(x_0, y_0; r; \phi)$.

We enumerate the sides of $p_n(x_0, y_0; r; \phi)$ in counter-clockwise fashion, $s_0, s_1, \cdots, s_{n-1}$, where s_0 is the side to which N is normal.

(5) J. L. Walsh considered finite averages over the vertices of regular n-gons, obtaining results similar to some of the results obtained in this paper, in his article *A mean-value theorem for polynomials and harmonic polynomials*, Bull. Amer. Math. Soc. vol. 42 (1936) pp. 923–930.

(6) Walsh, loc. cit. p. 924.

1.3. LEMMA 1. *If $f(x, y)$ is superficially summable in the interior of a finite domain D, if n is a fixed integer, $n \geqq 3$, and if for each point (x_0, y_0) in D the equation*

$$(8)\qquad f(x_0, y_0) = \frac{1}{|P_n(x_0, y_0; r; 0)|} \iint_{P_n(x_0, y_0; r; 0)} f(x_0 + \xi, y_0 + \eta) d\xi d\eta$$

holds for each $P_n(x_0, y_0; r; 0)$ in D, then $f(x, y)$ is harmonic in D(7).

Proof. For any function $F(x, y)$, defined and superficially summable in the interior of D, the areal averaging function

$$F(x, y; r) \equiv \frac{1}{|P_n(x, y; r; 0)|} \iint_{P_n(x, y; r; 0)} F(x + \xi, y + \eta) d\xi d\eta$$

is one degree smoother than $F(x, y)$ in the open subset D_r of D, where $F(x, y; r)$ is defined; that is, if $F(x, y)$ is superficially summable in the interior of D, then (for r fixed) $F(x, y; r)$ is (at least) continuous in D_r, or if $F(x, y)$ has continuous partial derivatives of the nth order in D, then $F(x, y; r)$ has continuous partial derivatives of the $(n+1)$st order in D_r(8). Hence it follows by a simple induction that if the summable function $f(x, y)$ satisfies (8) at each point (x_0, y_0) in D, for each $P_n(x_0, y_0; r; 0)$ lying in D, then $f(x, y)$ has continuous partial derivatives of all orders in D.

Accordingly, we may use the finite Taylor expansion

$$f(x_0 + \xi, y_0 + \eta) = f(x_0, y_0) + \sum_{k=1}^{2} \frac{1}{k!}\left[\left(\xi \frac{\partial}{\partial x} + \eta \frac{\partial}{\partial y}\right)^k f\right] + o(\zeta^2),$$

$$\zeta^2 = \xi^2 + \eta^2,$$

for $f(x, y)$, about the point (x_0, y_0); here

$$\xi \frac{\partial}{\partial x} + \eta \frac{\partial}{\partial y}$$

is a differential operator, the partial derivatives are evaluated at (x_0, y_0), and $o(r^\alpha)$ denotes a function (not always the same function) such that

$$\lim_{r\to 0} \frac{o(r^\alpha)}{r^\alpha} = 0.$$

The side s_m of $P_n(x_0, y_0; r; 0)$ can be represented by the polar equation

$$(9)\qquad s_m:\ \rho = r \sec(\theta - 2\pi m/n), \qquad (2m - 1)\pi/n \leqq \theta \leqq (2m + 1)\pi/n;$$

---

(7) In the proof we use only the weaker assumption that the difference of the two members in (8) is (uniformly) $o(r^4)$ in D.

(8) For a list of the principal properties of averaging functions, see H. E. Bray, *Proof of a formula for an area*, Bull. Amer. Math. Soc. vol. 29 (1923) pp. 264–270.

that is

$$s_m:\quad \rho = r\sec\psi,\qquad \psi = \theta - 2\pi m/n,\qquad -\pi/n \leqq \psi \leqq \pi/n.$$

Hence we have

$$\begin{aligned}\iint_{P_n(x_0,y_0;r;0)} f(x_0+\xi,\ y_0+\eta)d\xi d\eta &= \left| P_n(x_0,\ y_0;\ r;\ 0)\right| f(x_0,\ y_0)\\ &+ \sum_{m=0}^{n-1}\int_{-\pi/n}^{\pi/n}\int_0^{r\sec\psi}\sum_{k=1}^{2}\frac{1}{k!}\left\{\left[\rho\cos\left(\psi+\frac{2\pi m}{n}\right)\frac{\partial}{\partial x}\right.\right.\\ &\left.\left.+\rho\sin\left(\psi+\frac{2\pi m}{n}\right)\frac{\partial}{\partial y}\right]^k f(x_0,\ y_0)\right\}\rho d\rho d\psi + o(r^4).\end{aligned}$$

Applying (4), (5) and (6), we obtain

$$\begin{aligned}\iint_{P_n(x_0,y_0;r;0)} f(x_0+\xi,\ y_0+\eta)d\xi d\eta &= \left| P_n(x_0,\ y_0;\ r;\ 0)\right| f(x_0,\ y_0)\\ &+ \frac{nr^4}{8}\left(\frac{1}{3}\tan^3\frac{\pi}{n}+\tan\frac{\pi}{n}\right)\Delta f(x_0,\ y_0) + o(r^4),\end{aligned}\tag{10}$$

where

$$\Delta \equiv \frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}$$

is the Laplacian operator. The lemma now follows from (8) and (10), which yield the equation $\Delta f(x,\ y)=0$.

1.4. If $f(x,\ y)$ is harmonic in a finite domain D, then $f(x,\ y)$ may be expanded in a Fourier series about each point $(x_0,\ y_0)$ in D:

$$\begin{aligned}f(x,\ y) &= f(x_0+\rho\cos\theta,\ y_0+\rho\sin\theta)\\ &= f(x_0,\ y_0) + \sum_{k=1}^{\infty}\rho^k(a_k\cos k\theta + b_k\sin k\theta).\end{aligned}\tag{11}$$

LEMMA 2. *If $f(x,\ y)$ is harmonic in a finite domain D, then for each $P_n(x_0,\ y_0;\ r;\ 0)$ such that $D(x_0,\ y_0;\ r\sec\pi/n)$ is in D we have*

$$\begin{aligned}\iint_{P_n(x_0,y_0;r;0)} f(x_0+\xi,\ y_0+\eta)d\xi d\eta &= \left| P_n(x_0,\ y_0;\ r;\ 0)\right| f(x_0,\ y_0)\\ &+ n\sum_{k=1}^{\infty}\frac{a_{kn}r^{kn+2}}{kn+2}\int_{-\pi/n}^{\pi/n}\sec^{kn+2}\psi\cos kn\psi d\psi.\end{aligned}\tag{12}$$

Proof. Using (7), (9) and (11), we obtain

$$\iint_{P_n(x_0,y_0;r;0)} f(x_0+\xi,\, y_0+\eta)d\xi d\eta = |P_n(x_0,\, y_0;\, r;\, 0)|\, f(x_0,\, y_0)$$
$$+ n\int_{-\pi/n}^{\pi/n}\int_0^{r\sec\psi}\sum_{k=1}^{\infty}\rho^{kn}(a_{kn}\cos kn\psi + b_{kn}\sin kn\psi)\rho d\rho d\psi.$$

Now (12) follows from the fact that $\sin\theta$ is an odd function, while $\cos\theta$ and $\sec\theta$ are even functions, of θ.

2. Areal mean-values

2.1. The real and imaginary parts of $(x+iy)^n$ are basic homogeneous harmonic polynomials of degree n in the variables x, y. We shall denote these polynomials by $H_{1,n}(x, y)$ and $H_{2,n}(x, y)$, respectively. Any homogeneous harmonic polynomial of degree n in x, y is of the form

$$aH_{1,n}(x, y) + bH_{2,n}(x, y)$$

where a and b are constants.

2.2. Theorem 2. *If $f(x, y)$ is superficially summable in the interior of a finite domain D, and if n is a fixed integer, $n\geqq 3$, then a necessary and sufficient condition that for each point (x_0, y_0) in D, the equation*

$$f(x_0,\, y_0) = \frac{1}{|P_n(x_0,\, y_0;\, r;\, 0)|}\iint_{P_n(x_0,y_0;r;0)} f(x_0+\xi,\, y_0+\eta)d\xi d\eta \tag{13}$$

hold for each $P_n(x_0, y_0; r; 0)$ in D is that $f(x, y)$ be a harmonic polynomial of degree at most n, of the form

$$f(x, y) = B_0 + \sum_{k=1}^{n-1}[A_kH_{1,k}(x, y) + B_kH_{2,k}(x, y)] + B_nH_{2,n}(x, y), \tag{14}$$

where A_k and B_k are constants, $k=0, 1, \cdots, n$.

Necessity. If (13) holds, then, by Lemma 1, $f(x, y)$ is harmonic; consequently, by (13) and Lemma 2, we have

$$\sum_{k=1}^{\infty}\frac{a_{kn}r^{kn+2}}{kn+2}\int_{-\pi/n}^{\pi/n}\sec^{kn+2}\psi\cos kn\psi d\psi = 0. \tag{15}$$

Since (15) holds for all sufficiently small r, the coefficient of each power of r must vanish; in particular we have

$$\frac{a_n}{n+2}\int_{-\pi/n}^{\pi/n}\sec^{n+2}\psi\cos n\psi d\psi = 0. \tag{16}$$

Now

$$\int_{-\pi/n}^{\pi/n}\sec^{n+2}\psi\cos n\psi d\psi < 0,$$

as one readily sees by an inspection of the graphs of the functions sec ψ and cos $n\psi$. Therefore (16) yields

$$a_n = \frac{1}{n!} \frac{\partial^n}{\partial x^n} f(x_0, y_0) = 0.$$

But (13) holds at each point (x_0, y_0) in D, so that

$$\frac{\partial^n f(x, y)}{\partial x^n} = 0 \tag{17}$$

holds throughout D. By a simple induction we obtain, from (17) and the Laplace equation $\Delta f(x, y) = 0$,

$$\frac{\partial^{n+k} f(x, y)}{\partial x^s \partial y^t} = 0, \quad s + t = n + k;\, s, t = 0, 1, \cdots, n + k;\, k = 1, 2, \cdots. \tag{18}$$

Now (17) and (18) imply (14).

SUFFICIENCY. If $f(x, y)$ is of the form (14), then $f(x, y)$ can be continued harmonically so as to be defined and harmonic in the entire x, y-plane. Now (13) follows from (12), (17) and (18).

2.3. THEOREM 3. *If $f(x, y)$ is superficially summable in the interior of a finite domain D, if n is a fixed integer, $n \geqq 3$, and if ϕ_0 is fixed, $-\pi/n \leqq \phi_0 < \pi/n$, then a necessary and sufficient condition that for each point (x_0, y_0) in D, the equation*

$$f(x_0, y_0) = \frac{1}{|P_n(x_0, y_0; r; \phi_0)|} \iint_{P_n(x_0, y_0; r; \phi_0)} f(x_0 + \xi, y_0 + \eta) d\xi d\eta \tag{19}$$

hold for each $P_n(x_0, y_0; r; \phi_0)$ in D is that $f(x, y)$ be a harmonic polynomial of degree at most n, of the form

$$\begin{aligned} f(x, y) = B_0 + \sum_{k=1}^{n-1} [A_k H_{1,k}(x, y) + B_k H_{2,k}(x, y)] \\ + B_n [H_{1,n}(x, y) \sin n\phi_0 + H_{2,n}(x, y) \cos n\phi_0], \end{aligned} \tag{20}$$

where A_k and B_k are constants, $k = 0, 1, \cdots, n$.

Proof. If we make the transformation of coordinates $x' + iy' = (x + iy)e^{i\phi_0}$, then Theorem 3 follows at once from Theorem 2.

2.4. THEOREM 4. *If $f(x, y)$ is superficially summable in the interior of a finite domain D, and if n is a fixed integer, $n \geqq 3$, then a necessary and sufficient condition that for each point (x_0, y_0) in D, the equation*

$$f(x_0, y_0) = \frac{1}{|P_n(x_0, y_0; r; \phi)|} \iint_{P_n(x_0, y_0; r; \phi)} f(x_0 + \xi, y_0 + \eta) d\xi d\eta$$

hold for each $P_n(x_0, y_0; r; \phi)$ in D, is that $f(x, y)$ be a harmonic polynomial of degree at most $n-1$(9).

Proof. The theorem follows from Theorem 3 and the fact that while the class of harmonic polynomials of the form (20) is not invariant under rotations of the plane, the class of harmonic polynomials of degree less than n is invariant under these rotations.

3. Peripheral mean-values

3.1. Theorem 5. *If $f(x, y)$ is summable on each $P_n(x, y; r; 0)$ and on each $p_n(x, y; r; 0)$ lying in a finite simply-connected domain D, where n is a fixed integer, $n \geqq 3$, then a necessary and sufficient condition that for each point (x_0, y_0) in D the equation*

$$(21)\quad f(x_0, y_0) = \frac{1}{|\,p_n(x_0, y_0; r; 0)\,|} \int_{p_n(x_0, y_0; r; 0)} f(x_0 + \xi, y_0 + \eta) ds$$

hold for each $p_n(x_0\, y_0; r; 0)$ in D is that $f(x, y)$ be a harmonic polynomial of the form (14)(10).

Necessity. If we multiply both members of (21) by $|\,p_n(x_0, y_0; r; 0)\,|$ and integrate with respect to r, and then apply the theorem of Fubini to the superficially summable function $f(x, y)$, we obtain (13) for each point (x_0, y_0) in D, for each $P_n(x_0, y_0; r; 0)$ lying in D. Hence, by Theorem 2, $f(x, y)$ is of the form (14).

Sufficiency. If $f(x, y)$ is given by (14), then, by Theorem 2, (13) holds. Differentiating both sides of (13) with respect to r, we obtain

$$(22)\quad \begin{aligned} &\frac{1}{|\,p_n(x_0, y_0; r; 0)\,|} \int_{p_n(x_0, y_0; r; 0)} f(x_0 + \xi, y_0 + \eta) ds \\ &\qquad = \frac{1}{|\,P_n(x_0, y_0; r; 0)\,|} \iint_{P_n(x_0, y_0; r; 0)} f(x_0 + \xi, y_0 + \eta) d\xi d\eta, \end{aligned}$$

which, with (13), implies (21).

3.2. The following two theorems are analogous to Theorems 3 and 4, respectively.

Theorem 6. *If $f(x, y)$ is summable on each $P_n(x, y; r; \phi_0)$ and on each $p_n(x_0, y_0; r; \phi_0)$ lying in a finite simply-connected domain D, where n is a fixed integer, $n \geqq 3$, and where ϕ_0 is fixed, $-\pi/n \leqq \phi_0 < \pi/n$, then a necessary and sufficient condition that for each point (x_0, y_0) in D, the equation*

$$(23)\quad f(x_0, y_0) = \frac{1}{|\,p_n(x_0, y_0; r; \phi_0)\,|} \int_{p_n(x_0, y_0; r; \phi_0)} f(x_0 + \xi, y_0 + \eta) ds$$

(9) Cf. Walsh, loc. cit., p. 923, Theorem 3.

(10) Theorems 5–7 actually hold for $n=2$. We have stated them for $n \geqq 3$ to conform with the analogous Theorems 2–4, respectively.

hold for each $p_n(x_0, y_0, r; \phi_0)$ in D is that $f(x, y)$ be a harmonic polynomial of the form (20).

THEOREM 7. *If $f(x, y)$ is summable on each $P_n(x, y; r; \phi)$ and on each $p_n(x, y; r; \phi)$ lying in a finite simply-connected domain D, where n is a fixed integer, $n \geqq 3$, then a necessary and sufficient condition that for each point (x_0, y_0) in D, the equation*

$$f(x_0, y_0) = \frac{1}{|p_n(x_0, y_0; r; \phi)|} \int_{p_n(x_0, y_0; r; \phi)} f(x_0 + \xi, y_0 + \eta) ds$$

hold for each $p_n(x, y; r; \phi)$ in D is that $f(x, y)$ be a harmonic polynomial of degree at most $n-1$.

Theorems 6 and 7 follow from Theorems 3 and 4, respectively, in the same way that Theorem 5 follows from Theorem 2.

3.3. Theorems 5, 6 and 7 are analogous to Theorems 2, 3 and 4, respectively. Similarly, we might give three theorems of the type of Theorem 1 which are analogous to Theorems 2, 3 and 4. We give the explicit statement of only the last of these:

THEOREM 8. *If $f(x, y)$ is continuous in a finite simply-connected domain D, then a necessary and sufficient condition that for each point (x_0, y_0) in D, the equation*

$$\frac{1}{|p_n(x_0, y_0; r; \phi)|} \int_{p_n(x_0, y_0; r; \phi)} f(x_0 + \xi, y_0 + \eta) ds = \frac{1}{|P_n(x_0, y_0; r; \phi)|} \iint_{P_n(x_0, y_0; r; \phi)} f(x_0 + \xi, y_0 + \eta) d\xi d\eta$$

hold for each $P_n(x_0, y_0; r; \phi)$ lying in D, is that $f(x, y)$ be a harmonic polynomial of degree at most $n-1$.

THE UNIVERSITY OF MICHIGAN,
ANN ARBOR, MICH.
THE OHIO STATE UNIVERSITY,
COLUMBUS, OHIO

THE CONVERSE OF THE FATOU THEOREM FOR POSITIVE HARMONIC FUNCTIONS

BY

LYNN H. LOOMIS

1. **Introduction.** Let $v(z)=v(re^{i\phi})$ be a function harmonic in the unit circle $|z|<1$ and admitting there the Poisson-Stieltjes representation

$$v(re^{i\phi})=\frac{1}{2\pi}\int_0^{2\pi}\frac{1-r^2}{1+r^2-2r\cos(\theta-\phi)}\,dV(\theta), \tag{1}$$

where $V(\theta)$ is of bounded variation over $0\leqq\theta\leqq 2\pi$. The Fatou theorem([1]), in one form, has the following to say about the relation between $v(z)$ and $V(\theta)$ in (1):

THEOREM A. *If* $V_{(1)}(\theta)=\lim_{t\to 0}(V(\theta+t)-V(\theta-t))/2t$ *exists, then* $V(re^{i\theta})\to V_{(1)}(\theta)$ *as* $r\to 1$.

THEOREM B: *If the derivative* $V'(\theta)$ *exists, then* $v(z)\to V'(\theta)$ as $z\to e^{i\theta}$ *along any chord of* $|z|<1$ *(hence along any "non-tangential path" or "in angle")*.

The converses of these theorems are in general not true. If $v(z)$ is positive however, both converses can be proved. One result is that if $v(re^{i\theta})$ is a bounded function harmonic in $|z|<1$, and if its boundary function $v(\theta)$ is defined as the limit, wherever it exists, of $v(z)$ as $z\to e^{i\theta}$ "in angle," then $v(\theta)$ is a summable function which is *precisely* equal to the derivative of its indefinite integral. The converse of Theorem A for positive functions follows readily from known theorems, and it is the main object of this paper to deduce from it a strengthened form of the converse of Theorem B for positive functions.

We shall have occasion to use the theorem([2]) that a harmonic function has the representation (1) if and only if it can be written as the difference of two non-negative (or two positive) harmonic functions. In particular, every positive harmonic function has the representation (1) with $V(\theta)$ increasing.

2. **The converse of Theorem A for positive functions.** It will be simplest to infer the converse of Theorem A for positive functions from a series of remarks.

(i) The limit (if it exists) $V_{(1)}(\theta)=\lim_{t\to 0}[V(\theta+t)-V(\theta-t)]/2t$ is known as the generalized symmetric derivative of $V(\theta)$.

Presented to the Society, May 3, 1941 under the title *A converse to the Fatou theorem*; received by the editors May 12, 1942.

([1]) Fatou's original paper is in the Acta Math. vol. 30 (1906) pp. 335–400.

([2]) See Evans, *Logarithmic potential. Discontinuous Dirichlet and Neumann problems*, Amer. Math. Soc. Colloquium Publications vol. 6, 1927, p. 48.

(ii) If $\sum(a_n \cos n\theta + b_n \sin n\theta)$ is the Fourier-Stieltjes series for $dV(\theta)$, then $v(r^{i\theta}) = \sum r^n(a_n \cos n\theta + b_n \sin n\theta)$ is the Fourier series expansion for $v(z)$, so that the existence of the limit $\lim_{r\to 1} v(re^{i\theta})$ is equivalent (by definition) to the Abel summability (summability A) of the series $\sum(a_n \cos n\theta + b_n \sin n\theta)$.

(iii) It is a well known theorem[3] that a series is summable $(C, n+1)$ if it is summable A and if its nth Cesàro means are positive.

(iv) It is elementary that the $(C, 1)$ means of the Fourier-Stieltjes series of a non-decreasing function are positive[4].

(v) By a theorem of Hardy and Littlewood[5], summability (C, α) with $\alpha > 0$ for the Fourier-Stieltjes series of a non-decreasing function implies summability (C, β) for every $\beta > 0$, and is equivalent to the existence of the generalized symmetric derivative $V_{(1)}(\theta)$.

The converse of Theorem A for positive functions follows directly from these remarks.

3. **The Poisson-Stieltjes integral in the half-plane.** For some purposes it is convenient to work with integral representations in the half-plane rather than in the circle. A function $u(x, y)$ harmonic in the half-plane $y > 0$ admits the Poisson-Stieltjes integral representation

$$u(x, y) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{y(1 + t^2)}{y^2 + (t - x)^2} dU(t), \tag{2}$$

where $U(t)$ is of bounded variation over the closed infinite interval $[-\infty, \infty]$, if and only if the transformed function $v(w)$ obtained by mapping the half-plane (by $z = i\,(1-w)/(1+w)$) onto the unit circle $|w| < 1$ has the Poisson-Stieltjes representation (1), where $U(\tan \theta/2) = V(\theta)/2$. Note that (2) is not actually an improper integral, for the integrand is continuous over the closed infinite interval $[-\infty, \infty]$ and $U(t)$ is of bounded variation there; also note that $U(t)$ may have a jump at infinity. Obviously $U'(t) = V'(2 \text{ arc tan } t)/(1+t^2)$ when either derivative exists; thus $U'(0) = V'(0)$. We can rewrite (2) by removing the jump of $U(t)$ at infinity as ky and writing $U_1(t) = \int_0^t (1+t^2) dU(t)$. Then (2) becomes

$$u(x, y) = ky + \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{y}{y^2 + (t - x)^2} dU_1(t). \tag{3}$$

The integral is absolutely convergent and the kernel is simpler than the kernel of (2). Also $U_1'(t) = V'(2 \text{ arc tan } t)$ so that the Fatou theorem is generally valid. On the other hand, $U_1(t)$ is not of bounded variation in the in-

---

(3) See Kogbetliantz, *Sommation des séries et intégrales divergentes par les moyennes arithmétiques et typique*, Mémorial des Sciences Mathématiques vol. 51 p. 40, Theorem 21.

(4) See Titchmarsh, *Theory of functions*, p. 412. It is only necessary to replace the Lebesgue integral by a Stieljes integral in the equation for σ_n.

(5) See Zygmund, *Trigonometrical series*, p. 263 and p. 266, Example 11.

finite interval. We shall find it convenient to use (2) rather than (3), and adjust the mapping of $|w|<1$ onto $y>0$ so that any desired boundary point maps to the origin $z=0$ where the desired Fatou relation holds.

If $U_1(t)$ is absolutely continuous with derivative $u(t)$ then (3) becomes

$$u(x, y) = ky + \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{y}{y^2 + (t - x)^2} u(t)dt. \tag{4}$$

If the original function $v(w)$ admits the ordinary Poisson representation

$$v(re^{i\phi}) = \frac{1}{2\pi}\int_0^{2\pi} \frac{1 - r^2}{1 + r^2 - 2r\cos(\theta - \phi)} v(\theta)d\theta, \tag{5}$$

the transformed function has the representation (4) with $k=0$ and $u(t) = v(2 \arctan t)$.

For the purposes of this paper, the factor $1+t^2$ in the numerator of the integrand of (2) may be dropped. We are interested in limiting behavior as z approaches the origin along rays (rx, ry), $0<r\leqq 1$, $y>0$. But

$$\int_{-\infty}^{\infty} \frac{(ry)t^2}{(ry)^2 + (t - rx)^2} dU(t) = r\int_{-\infty}^{\infty} \frac{yt^2}{y^2 + (t - x)^2} dU(rt). \tag{6}$$

We consider an angle space by restricting x to $-x_0<x<x_0$. The integrand of the right member of (6) is bounded over $-x_0<x<x_0$, $-\infty\leqq t\leqq\infty$, so that the absolute value of the integral is bounded by rMV (where V is the variation of $U(t)$). Thus this term approaches 0 uniformly as z approaches the origin in any angle space, and we can disregard it. We have left to consider the harmonic function, again denoted $u(x, y)$,

$$u(x, y) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{y}{y^2 + (t - x)^2} dU(t), \tag{7}$$

which can be written

$$u(rx, ry) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{y}{y^2 + (t - x)^2} d\frac{U(rt)}{r}. \tag{8}$$

The Fatou Theorems A and B follow at once from (8). The assumption of Theorem A is that $[U(t)-U(-t)]/2t\to U_{(1)}(0)$ as $t\to 0$. Thus $[U(rt) - U(-rt)]/r = 2t(U)_{(1)}(0)+tR(rt)$ where $|R(t)|$ is bounded, say by M, and $|R(t)|\to 0$ as $t\to 0$. For $x=0$, (8) becomes

$$u(0, ry) = \frac{1}{\pi}\int_0^{\infty} \frac{y}{y^2 + t^2} d\left[\frac{U(rt) - U(-rt)}{r}\right], \tag{9}$$

and if we substitute the above expression in (9) and integrate the last term by parts, we have

$$
(10) \qquad
\begin{aligned}
u(0, ry) &= U_{(1)}(0) + \frac{2}{\pi}\int_0^\infty R(rt)\frac{t^2y}{(y^2+t^2)^2}\,dt \\
&= U_{(1)}(0) + \frac{2}{\pi}\int_0^{\pi/2} R(rt)\sin^2\beta d\beta,
\end{aligned}
$$

where $\beta = \arc\tan t/y$. For r small enough, $|R(rt)|$ is arbitrarily small over as large a part of the range of t, and hence of β, as desired. Over the remaining part of the range of β the integrand is bounded by M. Therefore the integral approaches 0, and $U_{(1)}(0) = \lim_{r\to 1} u(0, ry)$ which is the conclusion of Theorem A.

Theorem B can as easily be inferred. We assume $U(0)=0$, and have $U(rt)/r = tU'(0) + tR(rt)$, and instead of (10) this gives

$$
u(rx, ry) = U'(0) + \frac{2}{\pi}\int_{-\infty}^{\infty} R(rt)\frac{t(t-x)y}{[y^2+(t-x)^2]^2}\,dt.
$$

Now $|t(t-x)/(y^2+(t-x)^2)| < K$ over $-x_0 < x < x_0$, $-\infty \leqq t \leqq \infty$, and the absolute value of the integral is therefore bounded by

$$
\frac{2K}{\pi}\int_{-\pi/2}^{\pi/2} |R(rt)|\,d\beta,
$$

which approaches 0 with r as in the proof of Theorem A.

If one will compare these proofs with the corresponding proofs carried out in the unit circle[6], the advantages of the half-plane representations will be appreciated.

Using the representation (7) the converse of Theorem A for positive functions can be deduced immediately from the following integral Tauberian theorem of Hardy and Littlewood[7].

THEOREM. *Let $f(t)$ be positive, and suppose that $f(t)/(t+x)^\rho \in L(0, \infty)$ for some (and so for all) $x>0$. Suppose that*

$$
\int_0^\infty \frac{f(t)}{(t+x)^\rho}\,dt \sim \frac{H}{x^\sigma}
$$

as $x\to\infty$ (as $x\to 0$) for $0<\sigma<\rho$. Then

$$
F(t) = \int_0^t f(u)du \sim \frac{H\Gamma(\rho)}{\Gamma(\sigma)\Gamma(\rho-\sigma+1)}\,t^{\rho-\sigma}
$$

as $t\to\infty$ (as $t\to 0$).

The statement of the theorem can be modified to include Stieltjes integra-

(6) See Evans, loc. cit., pp. 39–43.

(7) Hardy and Littlewood, *On Tauberian theorems*, Proc. London Math. Soc. (2) vol. 30 (1930) p. 25.

tion, and then only elementary changes of variable are required to put the theorem in a form directly applicable to the Poisson integral for the half-plane.

We shall have occasion to use the following theorem(8).

THEOREM. *Let $U(t)$ have a jump m at $t=0$; thus $m=0$ is equivalent to the continuity of $U(t)$ at $t=0$. Then as $z=x+iy$ approaches the origin along the ray $x=Ry$, $yu(x, y)$ approaches the value $m/(1+k^2)\pi$. In particular, $U(t)$ is continuous at $t=0$ if and only if $yu(x, y)$ approaches 0 along some ray (and hence along all rays).*

In proof we consider

$$(ry)u(rx, ry) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{y^2}{y^2 + (t-x)^2}\, dU(rt).$$

Now as $r\to 0$, $U(rt)\to V(t)$ where $V(t)=U(0+)$ for $0<t<\infty$, $V(t)=U(0-)$ for $-\infty<t<0$, and $V(\pm\infty)=U(\pm\infty)$. Then by the Helly-Bray theorem concerning the convergence of sequences of Stieltjes integrals(9),

$$(ry)u(rx, ry) \to \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{y^2}{y^2 + (t-x)^2}\, dV(t) = \frac{my^2}{\pi(y^2 + x^2)}.$$

Since $x=ky$ the theorem follows.

4. **The converse of Theorem B for positive functions.** We shall prove the stronger theorem:

THEOREM 1. *Let $u(z)=u(x, y)$ be a positive harmonic function in the upper half-plane $y>0$, and having therefore the representation (2) with $U(t)$ increasing. If $\lim u(z)=u(0)$ as z approaches the origin along each of two rays, than $U'(0)$ exists and equals $u(0)$.*

We shall carry through the proof in a number of steps, using (7) instead of (2).

(i) If $u(z)\to u(0)$ as $z\to 0$ along each of two rays then $u(z)\to u(0)$ uniformly as $z\to 0$ between the rays. If the angle space between the rays is opened up to a half-plane by a power $w=e^{i\beta}z^{\alpha}$ we obtain from $u(z)$ a new positive harmonic function $u_1(x, y)$ continuous in the closed half-plane $y\geqq 0$ except possibly at the origin, and having a boundary function $u_1(t)$ which is continuous at the origin if it is defined there to have the value $u(0)$. Since $u_1(x, y)$ is positive it admits the representation (7) where $U_1(t)$ has the continuous derivative $u_1(t)$ when $t\neq 0$. Thus $U_1(t)$ is absolutely continuous if

(8) Fejér, *Über die Bestimmung des Sprunges der Funktion aus ihrer Fourierreihe*, J. reine angew. Math. vol. 142 (1913) pp. 165–166. See also Warschawski, *Bemerkung zu meiner Arbeit: Über das Randverhalten der Ableitung der Abbildungsfunktion bei konformer Abbildung*, Math. Zeit. vol. 38 (1934) p. 682.

(9) See D. V. Widder, *The Laplace transform*, p. 31.

it has no jump discontinuity at the origin. But by the theorem at the end of the last section, $U(t)$ is continuous at $t=0$, which implies successively that $ru(rx, ry)\to 0$ as $r\to 0$ for every point (x, y) in the upper half-plane $y>0$, that similarly $ru_1(rx, ry)\to 0$, and finally that $U_1(t)$ is continuous at the origin. Therefore $U_1(t)$ is absolutely continuous and $u_1(x, y)$ has the representation

$$u_1(x, y) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{y}{y^2 + (t - x)^2} u_1(t)dt.$$

The continuity of $u_1(t)$ at the origin implies the continuity of $u_1(x, y)$ at the origin, which proves the assertion (i).

(ii) Next, $u(z)\to u(0)$ as $z\to 0$ along any ray to the origin. Let l_1 and l_2 be the two given rays and suppose that l_2 makes the positive angle α with l_1. Let l_3 make the angle α with l_2 (we suppose that l_3 lies in the upper half-plane). We shall show that if $l:(rx, ry)$ is any ray between l_2 and l_3, then $u(rx, ry)\to u(0)$ as $r\to 0$. We need only to open up the angle space between l_1 and l_3 as in (i); l_2 becomes the ray perpendicular to the axis at the origin. The new positive harmonic function $u_1(x, y)$ has the representation (2) and by hypothesis (a) $u_1(0, ry)\to u(0)$ as $r\to 0$, (b) $U_1(t)$ is absolutely continuous for $t>0$, and $u_1(t)=U_1'(t)\to u(0)$ as $t\to 0$ (from the right). Thus $U_1(t)/t\to u(0)$ as $t\to 0$ from the right. Application of the converse of Theorem A for positive functions shows that $U_1(t)/t\to u(0)$ as $t\to 0$ from the left. Thus $U_1'(0)$ exists and equals $u(0)$. By the Fatou theorem $u_1(rx, ry)\to u(0)$ as $r\to 0$ for every (x, y) with $y>0$. This proves the statement about rays between l_2 and l_3. The assertion (i) together with a finite number of applications of the above process, proves (ii).

(iii) It remains to prove from these facts that $U'(0)$ exists and equals $u(0)$. By the converse of Theorem A it is sufficient to prove that $U(t)/t\to u(0)$ as $t\to 0$ from the right. The obvious device is to open up an angle space having the positive real axis as one of its bounding rays. But it is then somewhat difficult to establish the relation between the functions $U_1(t)$ and $U(t)$ for positive t. We shall proceed differently. Integrating (7) by parts we have

$$u(x, y) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{2y(t - x)}{[y^2 + (t - x)^2]^2} U(t)dt. \tag{11}$$

By (ii)

$$u(0, ry) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{2ty}{(y^2 + t^2)^2} \frac{U(rt)}{r} dt \to u(0) \qquad \text{as } r \to 0. \tag{12}$$

We can assume that $U(0)=0$, and the integrand in (12) is accordingly non-negative. We can therefore integrate and invert the order of integration, giving

$$\int_0^1 u(0, ry)dr = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{2ty}{(y^2+t^2)^2}\left(\int_0^1 \frac{U(rt)}{r} dr\right) dt$$
$$= \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{2ty}{(y^2+t^2)^2}\left(\int_0^t \frac{U(r)}{r} dr\right) dt.$$

In particular, $U(r)/r$ is integrable over every finite interval. We can now perform the same operation on (11), justifying the change in the order of integration by absolute integrability. Thus

$$u_1(x, y) = \int_0^1 u(rx, ry)dr = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{2y(t-x)}{[y^2+(t-x)^2]}\left(\int_0^t \frac{U(r)}{r} dr\right) dt,$$

and integrating by parts,

$$u_1(x, y) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{y}{y^2+(t-x)^2} \frac{U(t)}{t} dt.$$

The function $u_1(z) = u_1(x, y)$ is obviously harmonic and positive in the upper half-plane, and $u_1(z) \to u(0)$ as $z \to 0$ along any ray (rx, ry), $y > 0$. We now employ the device suggested at the beginning of (iii). Let l_1 and l_2 make angles α $(\alpha < \pi/2)$ and 2α with the positive real axis, and apply the transformation $w = z^{\pi/2\alpha}$. The harmonic function $u_1(z)$ has the boundary function $u_1(t) = U(t)/t$. After the transformation, the new harmonic function $u_2(w)$ has the boundary function $u_2(t) = u_1(t^{2\alpha/\pi})$. In the Stieltjes form $U_2'(t) = u_2(t)$ and by definition $U_2(t)/t \to u(0)$ as $t \to 0$ from the left. By the converse of Theorem A for positive functions

$$\frac{U_2(t)}{t} = \frac{1}{t}\int_0^t u_2(s)ds \to u(0)$$

as $t \to 0$ from the right. The following lemma is due to Landau([10]).

If $xf'(x)$ increases with x and $f(x) \sim x^a$ $(a > 0)$ as $x \to 0$, then $f'(x) \sim ax^{a-1}$ as $x \to 0$.

Here

$$f(t) = \int_0^t u_2(s)ds \sim tu(0)$$

as $t \to 0$, and $tf'(t) = tu_2(t) = t^{1-2\alpha/\pi}U(t^{2\alpha/\pi})$ which increases with t. Therefore by Landau's lemma, $u_2(t) \sim u(0)$ as $t \to 0$, that is, $U(t)/t \to u(0)$ as $t \to 0$ from the right. We now apply the converse of Theorem A again to obtain $U(t)/t \to u(0)$ as $t \to 0$ from the left. Thus $U'(0)$ exists and equals $u(0)$,and the proof of the theorem is complete.

([10]) E. Landau, *Beitrage zur analytischen Zahlentheorie*, Rend. Circ. Mat. Palermo vol. 24 (1917) pp. 81–160.

It should be remarked that the direct converse of Theorem B for positive functions can be proved from considerations of the integral representation (7) without any reference to the converse of Theorem A.

5. **A counterexample.** In this section we shall show by a counterexample that neither of the converses of A and B is true for the general representation (2). We first define the function $U(t)$ and then define the harmonic function $u(z)$ by the representation (7). The graph of $U(t)$ will consist of a sequence of triangular peaks separated by intervals of the t-axis and converging to the origin, the vertices of the peaks lying on the line $s=t$ over the points $t=2^{-n}$, the slopes of the sides of the peaks to be determined by later considerations. Such a function $U(t)$ is clearly of bounded variation.

We thus define $U(t)$ as follows:

$$U(t) = \begin{cases} 0, & t < 0 \\ 0, & t > 1 \\ (1/2)^n, & t = (1/2)^n, \quad n = 1, 2, \cdots \\ 0\,, & t = (1/2)^n \pm a_n, \quad 0 \leqq a_n \leqq 2^{-n-2}, \end{cases}$$

$U(t)$ linear on $2^{-n}-a_n \leqq t \leqq 2^{-n}$ and on $2^{-n} \leqq t \leqq 2^{-n}+a_n$, and $U(t)=0$ elsewhere in $0<t<1$. The a_n are positive numbers to be chosen later subject to the restriction noted above. On $2^{-n}-a_n<t<2^{-n}$, $dU(t)=2^{-n}dt/a_n$ so that

$$\begin{aligned} u(x, y) &= \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{y}{y^2+(t-x)^2}\, dU(t) \\ &= \frac{1}{\pi}\sum_{n=1}^{\infty}\left[\int_{2^{-n}-a_n}^{2^{-n}} \frac{1}{2^n a_n}\, \frac{y}{y^2+(t-x)^2}\, dt \right. \\ &\qquad\qquad \left. - \int_{2^{-n}}^{2^{-n}+a_n} \frac{1}{2^n a_n}\cdot\frac{y}{y^2+(t-x)^2}\, dt\right] \\ &= \frac{1}{\pi}\sum_{n=1}^{\infty}\int_{2^{-n}}^{2^{-n}+a_n} \frac{1}{2^n a_n}\left[\frac{y}{y^2+(t-x-a_n)^2} - \frac{y}{y^2+(t-x)^2}\right] dt \\ &= \frac{1}{\pi}\sum_{n=1}^{\infty}\int_{2^{-n}}^{2^{-n}+a_n} \frac{y}{2^n}\, \frac{2(t-x)-a_n}{[y^2+(t-x-a_n)^2][y^2+(t-x)^2]}\, dt. \end{aligned}$$

Thus

$$|u(x, y)| \leqq \frac{1}{\pi}\sum_{n=1}^{\infty}\frac{a_n}{2^n}\cdot \max_{2^{-n}\leqq t\leqq 2^{-n}+a_n}\left|\frac{y[2(t-x)-a_n]}{[y^2+(t-x-a_n)^2][y^2+(t-x)^2]}\right|.$$

Consider the term

$$\left|\frac{y}{2}\, \frac{2(t-x)-a_1}{[y^2+(t-x-a_1)^2][y^2+(t-x)^2]}\right|$$

on a ray $x=ky$. If we allow x, y, t and a_1 to vary, subject to the restrictions $0\leqq a_1\leqq 1/8$, $1/2-a_1\leqq t\leqq 1/2+a_1$, $x=ky$,the term has a maximum value M_k. By homogeneity the general term

$$\left|\frac{y}{2^n}\frac{2(t-x)-a_n}{[y^2+(t-x-a_n)^2][y^2+(t-x)^2]}\right|$$

with $0\leqq a_n\leqq 2^{-n-3}$, $2^{-n}-a_n\leqq t\leqq 2^{-n}+a_n$, $x=ky$ has the maximum value 2^nM_k. Now choose the constants a_n as 2^{-2n}. Then a_n times the general term above is bounded by $2^{-n}M_k$.

It is clear that the general term approaches 0 as y approaches 0 ($x=ky$) uniformly over the allowed range of t. Given ϵ, choose N so that

$$\sum_{n=N+1}^{\infty} 2^{-n}M_k < \epsilon/2,$$

and choose y_0 so that for $y<y_0$ and $x=ky$, the sum of the first N terms is bounded in absolute value by $\epsilon/2$. Thus for $y<y_0$ and $x=ky$, $|u(x, y)|<\epsilon$, and we have proved that $u(x, y)\to 0$ as $z\to 0$ along any ray to the origin. It is obvious however that $U(t)/t$ oscillates between 0 and 1 as $t\to 0$ from the right, and that $[U(t)-U(-t)]/2t$ oscillates between 0 and 1/2 as $t\to 0$ from the right.

6. **Generalizations and applications.** The procedure of Theorem 1 is adequate for situations more general than that described there. Suppose, for instance, we have not that $u(x, y)$ approaches $u(0)$ but that

$$\int_0^1 u(rx, ry)dr$$

exists and approaches $u(0)$ as $z=x+iy$ approaches the origin along each of two rays l_1 and l_2 We open up the angle space as before and get a positive harmonic function $u_1(x, y)$ admitting the representation

$$u_1(x, y)=\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{y}{y^2+(t-x)^2}u_1(t)dt$$

with the hypothesis that

$$\int_0^1 u_1(\pm|rt|^{\beta})dr\to u(0) \tag{13}$$

as $r\to 0$, two separate statements being understood. Here $\beta=\pi/\alpha$ where α is the angle between l_1 and l_2. We now need the following lemma:

LEMMA. *If $u(t)$ is a positive function such that for some $\alpha>0$*

$$\int_0^s u(t^{\alpha})dt \sim ks$$

as $s \to 0$, *then*

$$\int_0^s u(t)dt \sim ks. \tag{14}$$

In proof we multiply both sides by $s^{\alpha-2}$ and integrate from 0 to r. If we then integrate the left member by parts the conclusion follows.

If rt is replaced by t in (13) and the lemma applied, we obtain as a conclusion precisely the hypothesis of the Fatou Theorem B for the harmonic function $u_1(x, y)$ (with $U(s)$ absolutely continuous and equal to the integral (14)). Therefore $u_1(z) \to u(0)$ as $z \to 0$ "in angle," which is equivalent to the statement that $u(z) \to u(0)$ as $z \to 0$ along any path between l_1 and l_2. We can therefore apply Theorem 1 to infer that $U'(0)$ exists and equals $u(0)$. Also

$$\int_0^1 u(sx, sy)ds = \frac{1}{r}\int_0^r u(t\cos\theta, t\sin\theta)dt$$

where $z = x+iy = r(\cos\theta + i\sin\theta)$, and the new hypothesis is thus that the integral Hölder mean approaches $u(0)$. We have thus proved the following theorem:

THEOREM 2. *Let* $u(z) = u(x, y)$ *be a positive harmonic function in the upper half-plane* $y > 0$, *and having therefore the representation* (2) *with* $U(t)$ *increasing. If* $u(z)$ *has the* $(H, 1)$ *limit* $u(0)$ *as* $z = x+iy$ *approaches the origin along each of two rays, then* $U'(0)$ *exists and equals* $u(0)$.

COROLLARY 1. *As a consequence of the Fatou theorem* $u(x, y)$ *has the ordinary limit* $u(0)$ *as* $z = x+iy$ *approaches the origin along any ray of the upper half-plane.*

COROLLARY 2. *If* $u(z)$ *has the* (H, n) *limit* $u(0)$ *as* z *approaches the origin along each of two rays, then* $U'(0)$ *exists and equals* $u(0)$.

This is a trivial consequence of the Landau lemma used in §4, for if $u(t)$ is positive, and

$$\int_0^s \frac{1}{t}\int_0^t u(r)dr \sim su(0) \qquad s \to 0,$$

then direct application of the lemma gives that

$$\int_0^t u(r)dr \sim tu(0).$$

A finite number of such steps reduces the hypothesis of the corollary to that of Theorem 2.

Again, it is clear that we have used the monotonicity of $U(t)$ (positivity

of $u(x, y)$) only locally about the origin, and the hypothesis can be accordingly weakened to that extent.

It is interesting to see what can be said for other statements of the Fatou theorem. A somewhat stronger form of the theorem than that contained in the first section concerns the ordinary Poisson integral representation

$$v(re^{i\phi}) = \frac{1}{2\pi}\int_0^{2\pi} \frac{1 - r^2}{1 + r^2 - 2r\cos(\theta - \phi)} v(\theta)d\theta.$$

Let $v_\theta(z) = \partial v(re^{i\theta})/\partial\theta$. Then $v_\theta(z)$ is a harmonic function in $|z| < 1$ which does not in general admit a Poisson-Stieltjes integral representation.

THEOREM A′. *If* $v_{(1)}(\theta) = \lim_{t\to 0} [v(\theta+t) - v(\theta-t)]/2t$ *exists, then* $v_\theta(re^{i\theta}) \to v_{(1)}(\theta)$ *as* $r \to 1$.

THEOREM B′. *If the derivative* $v'(\theta)$ *exists, then* $v_\theta(z) \to v'(\theta)$ *uniformly as* $z \to e^{i\theta}$ *"in angle."*

The proofs are essentially the same as those for Theorems A and B. As before, we must impose some further restriction on $v(\theta)$ in order to deduce the converses of Theorems A′ and B′, and we try the local condition that, $v(\theta) - v(\theta_0)$ change sign at θ_0 that is, that $[v(\theta) - v(\theta_0)](\theta - \theta_0)$ be of constant sign (admitting the value 0) in some neighborhood of θ_0. We may obviously take $v(\theta_0) = 0$. Thus in the half-plane our hypotheses are that $u(x, y)$ has the representation

$$u(x, y) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{y}{y^2 + (t - x)^2} u(t)dt,$$

where the integral is absolutely convergent, that $tu(t) \geqq 0$ in some neighborhood of the origin, and the $\partial u(x, y)/\partial x = u_x(x, y)$ has the property that $u_x(rx, ry) \to l$ as $r \to 0$ for every (x, y) with $y > 0$. For the converse of A, the assumption holds only along the ray $x = 0$. Now

$$u_x(x, y) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{2y(t - x)}{[y^2 + (t - x)^2]} u(t)dt,$$

and

$$u_x(0, ry) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{2yt}{[y^2 + t^2]^2} \frac{u(rt)}{r} dt.$$

The integrand here is positive in the neighborhood of $r = 0$, $t = 0$, and since $u_x(0, ry)$ is integrable over $0 < r < 1$ we have

$$\int_0^1 u_x(0, ry)dr = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{2yt}{[y^2 + t^2]^2}\left(\int_0^t \frac{u(r)}{r} dr\right)dt,$$

and in general

$$\int_0^1 u_x(rx, ry)dr = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{2y(t-x)}{[y^2+(t-x)^2]}\left(\int_0^t \frac{u(r)}{r} dr\right)dt$$
$$= \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{y}{y^2+(t-x)^2}\,\frac{u(t)}{t}\, dt.$$

Since $u(t)/t$ is non-negative in some neighborhood of $t=0$, we can apply Theorem 1 with $u(t)/t=U'(t)$ to obtain

$$\text{(a)} \qquad \frac{1}{t}\int_0^t \frac{u(r)-u(-r)}{2r}\, dr \to l$$

as $t \to 0$, as the converse of Theorem A′, and

$$\text{(b)} \qquad \frac{1}{t}\int_0^t \frac{u(r)}{r}\, dr \to l$$

as $t \to 0$, as the converse of Theorem B′.

To obtain the symmetry of the earlier case we should now prove that (a) and (b) can be taken as weakened hypotheses for the Fatou Theorems A′ and B′. This is in fact the case, but we shall omit the proofs here since they are essentially the same as the proofs of Theorems A and B.

The relation between $v_\theta(w)$ in the unit circle and $u_x(x, y)$ in the half-plane can be easily established. They are different functions even when transformed so as to have the same domain of definition, but they have the same asymptotic properties at the origin in the half-plane.

If $f(z)$ is a bounded analytic function in the unit circle $|z|<1$ it is, known[11] that if $\lim f(z)$ exists as z approaches a boundary point $e^{i\theta}$ along some curve, then $f(z)$ has that limit as $z \to e^{i\theta}$ "in angle." Thus there is no difference between situations A and B in this case. The Fatou theorem implies[12] that $\lim_{r\to 1} f(re^{i\theta})$ exists for almost all θ, and the converse of the Fatou theorem (Theorem 1) implies that if $f(\theta)$ is the boundary function thus defined then $f(\theta)$ is *precisely* equal to the derivative of its indefinite integral.

---

(11) See Nevanlinna, *Eindeutige Analytische Funktionen*, p. 65.
(12) See Bieberbach, *Funktionentheorie*, II, pp. 147–148.

Harvard University,
Cambridge, Mass.

ON BOUNDED VARIATION AND ABSOLUTE CONTINUITY FOR PARAMETRIC REPRESENTATIONS OF CONTINUOUS SURFACES

BY

PAUL V. REICHELDERFER

INTRODUCTION

1. A continuous curve C in xyz-space may be defined by

$$C:\quad x = x(u),\quad y = y(u),\quad z = z(u),\quad a \leqq u \leqq b,$$

where each of the functions $x(u), y(u), z(u)$ is continuous on the closed interval $[a, b]$. The following facts are known[1] (see Rado [3, chap. I]; Saks [1, chap. IV]).

(1) A necessary and sufficient condition that the length $L(C)$ of C be finite is that each of the functions $x(u), y(u), z(u)$ be of bounded variation on $[a, b]$.

(2) If the length $L(C)$ is finite, then each of the derivatives $x'(u), y'(u), z'(u)$ exists almost everywhere in $[a, b]$, is summable on $[a, b]$, and

$$L(C) \geqq \int_a^b [x'(u)^2 + y'(u)^2 + z'(u)^2]^{1/2} du.$$

(3) A necessary and sufficient condition that the sign of equality hold in this relation is that each of the functions $x(u), y(u), z(u)$ be absolutely continuous on $[a, b]$.

2. A continuous surface S in xyz-space may be defined by

$$S:\quad x = x(u, v),\quad y = y(u, v),\quad z = z(u, v),\quad a \leqq u \leqq b,\quad c \leqq v \leqq d,$$

where each of the functions $x(u, v), y(u, v), z(u, v)$, is continuous on the closed (two-dimensional) interval $[a, b; c, d]$. How may the concepts for the area of the surface S and for bounded variation and absolute continuity of the representation of S be defined so that theorems analogous to those for continuous curves cited in[2] 1 hold? For the special case in which S may be given by relations of the form

$$S:\quad x = u,\quad y = v,\quad z = f(u, v),\quad a \leqq u \leqq b,\quad c \leqq v \leqq d,$$

Presented to the Society, April 17, 1942; received by the editors June 11, 1942.

[1] Numbers in square brackets refer to papers listed in the bibliography at the end of this paper.

[2] The notation I, 6, 2, for example, refers to chapter I, section 6, relation 2 in this paper When no chapter reference is given, the introduction is meant.

Geöcze and Tonelli have shown a complete answer to this question (see Geöcze [1], Tonelli [1]). But for the general case, no satisfactory answer seems to be known.

3. It is the chief purpose of this paper to give an answer to the question just raised. In so doing, interesting generalizations and extensions of results in the literature will be obtained. Briefly, the program for procedure is the following. First, a definition for a continuous surface is made precise (see I, 1). Now concepts for bounded variation and absolute continuity of a representation for a curve are phrased in terms of the corresponding representations for the projections of this curve upon each of the coordinate axes. Here the definitions for bounded variation and absolute continuity of a representation for a surface will be made in terms of the corresponding representations for the projections of this surface upon each of the coordinate planes. For representations of the latter type, a hierarchy of definitions for bounded variation and absolute continuity is extant(3) (see R² [1]). It will be desirable for the purpose of this paper to review these definitions, and to make certain additions to the theory developed in the work just cited (see II, III). Next, a definition for the area of a continuous surface will be given (see IV). This definition will be compared with that of the Geöcze area as defined by Rado (see IV, **17–20**), and the Lebesgue area (see IV, **14–15**); the latter area has been most frequently used in the literature. For the special case considered by Geöcze and Tonelli, it will be shown that the definitions advanced here are equivalent to those which they used (see IV, **16**). With the definitions for the area of a surface, and for bounded variation and absolute continuity of its representations, thus formulated, it will be shown that theorems hold for continuous surfaces which are analogous to those cited in **1** for curves (see IV, **4**, **6–13**; V, **9–15**). As an application of this theory, some of the results of Rado and Reichelderfer on convergence in area for surfaces (see R² [2]) will be generalized (see V).

4. For brevity, the following notations and conventions are adopted. The u^1u^2-plane will serve as a parameter plane; a point (u^1,u^2) in it is denoted simply by u. The surfaces will be in $x^1x^2x^3$-space; a point (x^1, x^2, x^3) in this space is denoted simply by x. With each point x, there is associated its projections on the respective coordinate planes given by

$$^1x = (0, x^2, x^3), \qquad ^2x = (x^1, 0, x^3), \qquad ^3x = (x^1, x^2, 0).$$

The planar exterior measure of any set E in the u-plane is denoted by $|E|$. The set of interior points in the set E is denoted by E^0.

If $a=(a^1, a^2, a^3)$ be any triple of real numbers, set

$$|a| = (|a^1|, |a^2|, |a^3|), \qquad \|a\| = [(a^1)^2 + (a^2)^2 + (a^3)^2]^{1/2}.$$

A triple of real, finite, single-valued functions $x^i(u)$, $i=1, 2, 3$, each defined

(3) The symbol R² in this note is to be read "Rado and Reichelderfer."

on a set E in the u-plane is denoted by $[x(u), E]$, where $x(u)=(x^1(u), x^2(u), x^3(u))$ for $u=(u^1, u^2)$ in E. With each triple $[x(u), E]$ there are associated three triples $[{}^ix(u), E]$ defined by

$$^1x(u) = (0, x^2(u), x^3(u)),\quad {}^2x(u) = (x^1(u), 0, x^3(u)),\quad {}^3x(u) = (x^1(u), x^2(u), 0),\qquad u \text{ in } E.$$

A triple $[x(u), E]$ is said to possess any property which is possessed by each of the $x^i(u)$ for $i=1, 2, 3$ on the set E.

A two-dimensional interval in the u-plane is denoted generically by I, $\mathfrak{I}$, $[\alpha, \beta]$, or $[\alpha^1, \beta^1; \alpha^2, \beta^2]$; it consists of all points u satisfying $\alpha^1 \leqq u^1 \leqq \beta^1$, $\alpha^2 \leqq u^2 \leqq \beta^2$ where $\alpha^1 < \beta^1$, $\alpha^2 < \beta^2$. A connected open set in the u-plane is termed a domain, is denoted generically by $\mathcal{D}$ or $\mathfrak{D}$; if the boundary of a domain consists of a Jordan curve, then the closed connected set of points in the domain and on its boundary is termed a simple Jordan region, is denoted generically by β, B or $\mathfrak{B}$. If the boundary of a domain consists of a finite number of Jordan curves, then the closed connected set of points in the domain and on its boundary is termed a Jordan region, is denoted generically by $\mathfrak{R}$. A sequence of domains $\mathcal{D}_n$ is said to fill up a domain $\mathcal{D}$ from the interior if each domain $\mathcal{D}_n$ is contained in $\mathcal{D}$, but for every closed set F in $\mathcal{D}$ there exists an $n(F)$ such that F is in $\mathcal{D}_n$ for every choice of n exceeding $n(F)$. A sequence of Jordan regions $\mathfrak{R}_n$ fill up a Jordan region $\mathfrak{R}$ from the interior if their interiors $\mathfrak{R}_n^0$ fill up $\mathfrak{R}^0$ from the interior.

If $\mathfrak{B}$ is any simple Jordan region in the u-plane, then a finite system of nonoverlapping simple Jordan regions B lying in $\mathfrak{B}$ is denoted generically by $S(\mathfrak{B})$. The maximum of the diameters of the domains B in $S(\mathfrak{B})$ is denoted by $\|S(\mathfrak{B})\|$. If $\mathfrak{B}=\sum B$ for B in $S(\mathfrak{B})$, then $S(\mathfrak{B})$ is termed a subdivision of $\mathfrak{B}$. In particular, if each of the simple Jordan regions B in $S(\mathfrak{B})$ is an interval, then $S(\mathfrak{B})$ is termed a finite interval system; if, moreover, $\mathfrak{B}$ is an interval and $S(\mathfrak{B})$ is a subdivision of $\mathfrak{B}$, then $S(\mathfrak{B})$ is termed an interval subdivision.

A b-function defined in $\mathfrak{B}$ is a law which associates with every simple Jordan region B in $\mathfrak{B}$ a finite, real number $\phi(B)$; this function is denoted by $[\phi, \mathfrak{B}]$. This b-function is non-negative if $\phi(B)$ is non-negative for every B in $(\mathfrak{B})$. For a finite system $S(B)$, where B is any simple Jordan region in $\mathfrak{B}$, set

$$\phi(S(B)) = \sum \phi(\beta) \quad \text{for} \quad \beta \text{ in } S(B);$$
$$U(B; [\phi, \mathfrak{B}]) = \text{l.u.b. } \phi(S(B)) \quad \text{for all finite systems} \quad S(B).$$

Evidently $\phi(B) \leqq U(B; [\phi, \mathfrak{B}]) \leqq U(\mathfrak{B}; [\phi, \mathfrak{B}])$ for every choice of B in $\mathfrak{B}$. Hence if $U(\mathfrak{B}; [\phi, \mathfrak{B}])$ is finite, then $[U, \mathfrak{B}]$ is a b-function, and $[\phi, \mathfrak{B}]$ is said to possess a U-function. If $S(\mathfrak{B})$ is any finite system, then clearly $U(S(\mathfrak{B}); [\phi, \mathfrak{B}]) \leqq U(\mathfrak{B}; [\phi, \mathfrak{B}])$.

5. If $[{}^i\phi, \mathfrak{B}]$, $i=1, 2, 3$, is a triple of b-functions having a common range of definition, set

$$\phi(B) = ({}^1\phi(B), {}^2\phi(B), {}^3\phi(B)), \qquad \Phi(B) = \|\phi(B)\|, \quad \text{for} \quad B \text{ in } \mathfrak{B}.$$

If $[\phi, \mathfrak{B}]$ is a triple of non-negative b-functions, then clearly

$${}^i\phi(B) \leqq \Phi(B) \leqq \sum_{i=1}^{3} {}^i\phi(B) \quad \text{for} \quad B \text{ in } \mathfrak{B}.$$

Hence a necessary and sufficient condition that $[\Phi, \mathfrak{B}]$ have a U-function is that each member of the non-negative triple $[\phi, \mathfrak{B}]$ have a U-function, and

1. $$U(B; [{}^i\phi, \mathfrak{B}]) \leqq U(B; [\Phi, \mathfrak{B}]) \leqq \sum_{i=1}^{3} U(B; [{}^i\phi, \mathfrak{B}]) \quad \text{for} \quad B \text{ in } \mathfrak{B}.$$

Elementary considerations lead to the following

LEMMA. *Let* $[\phi_n, \mathfrak{B}]$, $n=0, 1, 2, \cdots$, *be a sequence of triples of non-negative b-functions for which* $\liminf {}^i\phi_n(B) \geqq {}^i\phi_0(B)$ *for* B *in* $\mathfrak{B}$, $i=1, 2, 3$. *Then* $\liminf \Phi_n(B) \geqq \Phi_0(B)$; $\liminf U(B; [{}^i\phi_n, \mathfrak{B}]) \geqq U(B; [{}^i\phi_0, \mathfrak{B}])$; $\liminf U(B; [\Phi_n, \mathfrak{B}] \geqq U(B; [\Phi_0, \mathfrak{B}])$ *for* B *in* $\mathfrak{B}$.

CHAPTER I

ON CONTINUOUS SURFACES

1. A definition for a continuous surface will now be recalled; since this definition is in the literature (see Rado [2, 3]), it will be merely sketched here for the convenience of the reader, and for the purpose of fixing notation in the sequel. Consider the class of all continuous triples $[x(u), \mathfrak{B}]$, where $\mathfrak{B}$ is a simple Jordan region in the u-plane. Let $[x_1(u), \mathfrak{B}_1]$, $[x_2(u), \mathfrak{B}_2]$ be any two of these triples. Since $\mathfrak{B}_1$ and $\mathfrak{B}_2$ are simple Jordan regions, there exist topological maps of $\mathfrak{B}_1$ onto $\mathfrak{B}_2$ given by single-valued continuous pairs $[\bar{u}(u), \mathfrak{B}_1]$ having single-valued continuous inverses on $\mathfrak{B}_2$. Let $d(\bar{u})$ denote the maximum of $\|x_1(u) - x_2(\bar{u}(u))\|$ for u in $\mathfrak{B}_1$. Put $d([x_1, \mathfrak{B}_1], [x_2, \mathfrak{B}_2])$ equal to the greatest lower bound of $d(\bar{u})$ for all topological maps $[\bar{u}(u), \mathfrak{B}_1]$ of $\mathfrak{B}_1$ onto $\mathfrak{B}_2$. It is easily verified that the binary relation d thus defined in the class of all continuous triples has all the properties of a distance except one: the fact that $d([x_1, \mathfrak{B}_1], [x_2, \mathfrak{B}_2])$ is zero does not imply that $\mathfrak{B}_1$ and $\mathfrak{B}_2$ are identical and $x_1(u) = x_2(u)$ for u in $\mathfrak{B}_1 \cdot \mathfrak{B}_2$. In order to remedy this defect, one agrees that two of these triples $[x_1, \mathfrak{B}_1]$, $[x_2, \mathfrak{B}_2]$ are in the $\sim$ relation provided $d([x_1, \mathfrak{B}_1], [x_2, \mathfrak{B}_2]) = 0$. It is readily verified that the binary relation $\sim$ is an equivalence relation; hence it partitions the class of all continuous triples into mutually exclusive sets of triples mutually in the $\sim$ relation; denote these sets generically by S. It follows that if S_1 and S_2 are any two of these sets, then $d([x_1, \mathfrak{B}_1], [x_2, \mathfrak{B}_2])$ has a value d_{12} which is independent of the choice of $[x_1, \mathfrak{B}_1]$ in S_1 and $[x_2, \mathfrak{B}_2]$ in S_2. Set $d(S_1, S_2) = d_{12}$. Then $d(S_1, S_2)$ has all the properties of a distance in the class of sets S. Each of the sets S is termed a *continuous surface* of the type of the circular disc. Any one

of the continuous triples $[x(u), \mathfrak{B}]$ in S is termed a (parametric) *representation* for the surface S. The distance $d(S_1, S_2)$ is known as the *Fréchet distance* of the surfaces S_1 and S_2. A sequence of surfaces S_n is said to converge to the surface S_0 if $d(S_n, S_0)$ converges to zero. If S_n converges to S_0, and if $[x_0(u), \mathfrak{B}_0]$ be any representation for S_0, then there exist representations $[x_n(u), \mathfrak{B}_n]$ for S_n such that $\mathfrak{B}_n$ is identical with $\mathfrak{B}_0$ for every n and $x_n(u)$ converges on $\mathfrak{B}_0$ uniformly to $x_0(u)$.

2. Amongst the representations for a surface S, there may occur one $[x(u), \mathfrak{B}]$ of the form

$$1. \qquad x(u) = (u^1, u^2, x^3(u^1, u^2)), \qquad u = (u^1, u^2) \text{ in } \mathfrak{B}.$$

Denote by ${}^3\mathfrak{B}$ the image of $\mathfrak{B}$ in the ${}^3x$-plane under the topological map $x^1 = u^1$, $x^2 = u^2$ for (u^1, u^2) in $\mathfrak{B}$. Then

$$x^3 = x^3(x^1, x^2), \qquad (x^1, x^2) \text{ in } {}^3\mathfrak{B}.$$

denotes what is commonly called a non-parametric representation for S. For this reason, the representation $[x(u), \mathfrak{B}]$ in 1 will be termed a *representation of non-parametric origin* for S. By symmetry, one should also term any representation for S having one of the forms

$$(u^2, u^1, x^3(u^1, u^2)), \qquad (u^1, x^2(u^1, u^2), u^2), \qquad (u^2, x^2(u^1, u^2), u^1),$$
$$(x^1(u^1, u^2), u^1, u^2), \qquad (x^1(u^1, u^2), u^2, u^1),$$

a representation of non-parametric origin for S. But this is unnecessary, since any of these forms may be brought into form 1 by a suitable change of notation. A surface need not have a representation of non-parametric origin, but if it does, that representation is unique.

3. Amongst the representations for a surface S, there may occur one $[x(u), \mathfrak{B}]$ of the form

$$1. \qquad x(u) = (x^1(u^1, u^2), x^2(u^1, u^2), 0), \qquad u = (u^1, u^2) \text{ in } \mathfrak{B}.$$

Geometrically speaking, such a surface lies entirely in the ${}^3x$-plane. It is easily verified that any other representation for S must have the form 1. Such a surface is sometimes called a *flat surface*. The representation for S defines a *continuous transformation* from the simple Jordan region $\mathfrak{B}$ in the u-plane to a bounded portion of the ${}^3x$-plane.

4. Let S be any continuous surface in x-space. If $[x(u), \mathfrak{B}]$ be any representation for S, then (see 4) $[{}^3x(u), \mathfrak{B}]$ is a representation for a flat surface ${}^3S$ in the ${}^3x$-plane, which is the projection of S on that plane. It follows at once (see I, 1) that if $[\bar{x}(u), \overline{\mathfrak{B}}]$ is any other representation for S then $[{}^3\bar{x}(u), \overline{\mathfrak{B}}]$ is another representation for ${}^3S$—that is, the surface ${}^3S$ is uniquely determined by S. Thus a continuous surface S in x-space determines uniquely three flat projection surfaces ${}^1S$, ${}^2S$, ${}^3S$ on the coordinate planes ${}^1x$, ${}^2x$, ${}^3x$, respectively.

5. A real, finite-valued function $f(u)=f(u^1, u^2)$ defined on a simple polygonal region $\mathfrak{B}$ in the u-plane is termed *quasi-linear* if $f(u)$ is continuous in $\mathfrak{B}$, and if there exists a triangulation of $\mathfrak{B}$ such that $f(u)$ is a linear function of u^1 and u^2 on each triangle of the triangulation. A continuous surface is termed a polyhedron, and denoted by P, if it possesses a representation $[x(u), \mathfrak{B}]$ such that $\mathfrak{B}$ is a simple polygonal region and each $x^i(u)$ is quasi-linear on $\mathfrak{B}$. Then there exists a triangulation of $\mathfrak{B}$ such that each of the functions $x^i(u)$ for $i=1, 2, 3$ is linear on every triangle in the triangulation. The image of each triangle in this triangulation is a (possibly degenerate) triangle; the sum of the areas of these image triangles is termed the elementary area of P—denote it by $a(P)$ Elementary considerations show that $a(P)$ depends only on the polyhedron P.

6. Let S be any continuous surface. Then there always exist sequences of polyhedra P_n such that P_n converges to S; lim inf $a(P_n)$ is an upper bound for the *Lebesgue area* $A(S)$ of the surface; $A(S)$ is the greatest lower bound of all the upper bounds derived in this way. The Lebesgue area possesses the following important properties (see Rado [2, 3]).

1. If P be any polyhedron, then $A(P)=a(P)$.
2. There exists a sequence of polyhedra P_n such that P_n converges to S and $A(P_n)$ converges to $A(S)$.
3. The Lebesgue area $A(S)$ is a lower semi-continuous functional of S.

7. Much of the literature on continuous surfaces restricts its considerations to surfaces having representations of non-parametric origin (see I, **2**); in defining the "Lebesgue area" for such surfaces, it has been convenient to restrict the class of approximating polyhedra P_n to have representations of non-parametric origin also. Let $A_*(S)$ denote the area of a surface S having a representation of non-parametric origin when the class of approximating polyhedra is so restricted. Then clearly $A(S) \leqq A_*(S)$, and it is important in comparing the literature to know that the sign of equality always holds. This fact is implicit in the work of Rado (see Rado [1]), but no explicit proof seems to be in the literature. Such a proof will be a corollary to one of the results in this paper (see IV, **15**).

8. Let S be any continuous surface. The following principle has been advanced by Rado and Reichelderfer to direct their work in the theory of continuous surfaces (see R² [2]). Assume that some sort of area—denote it by $\mathcal{A}(S)$—is defined for S and that, for each representation $[x(u), \mathfrak{B}]$ of S, some sort of Jacobians are defined for each of the projection representations $[{}^1x, \mathfrak{B}]$, $[{}^2x, \mathfrak{B}]$, $[{}^3x, \mathfrak{B}]$ on the coordinate planes (see **4**)—denote these by ${}^1\mathcal{J}(u)$, ${}^2\mathcal{J}(u)$, ${}^3\mathcal{J}(u)$, respectively, wherever they exist. Then a representation $[x(u), \mathfrak{B}]$ for S is said to be *absolutely continuous* ($\mathcal{A}$, $\mathcal{J}$) provided that each of the Jacobians ${}^1\mathcal{J}(u)$, ${}^2\mathcal{J}(u)$, ${}^3\mathcal{J}(u)$ exists almost everywhere in the interior $\mathfrak{B}^0$ of $\mathfrak{B}$, is summable on $\mathfrak{B}^0$, and

$$\mathcal{A}(S) = \int_{\mathfrak{B}^*} [{}^1\mathcal{J}(u)^2 + {}^2\mathcal{J}(u)^2 + {}^3\mathcal{J}(u)^2]^{1/2} du.$$

Chapter II

On continuous transformations in the plane

In sections **1–9**, **13**, the salient features of the theory of bounded variation and absolute continuity for continuous transformations in the plane developed by Rado and Reichelderfer (see R² [1]) are summarized. For all details, the reader is referred to the cited paper(4). Minor notational changes have been made to place the results in a form more convenient for the purposes of this paper. In sections **10–12**, **14–22** extensions of this theory are discussed.

1. Let ξ be any plane in x-space; on ξ choose a rectangular coordinate system ξ^1, ξ^2, and adopt notations similar to those introduced in **4** for the u-plane. Let $\mathfrak{D}$ be any bounded domain in the u-plane. If $\xi(u) = (\xi^1(u^1, u^2), \xi^2(u^1, u^2))$ be a pair of real, single-valued functions defined, continuous, and bounded in $\mathfrak{D}$, then $[\xi(u), \mathfrak{D}]$ defines a bounded continuous transformation T, which associates with every point u in $\mathfrak{D}$ a point $\xi = \xi(u)$ in a bounded portion of the ξ-plane. If E be any set in the u-plane, let $T(E)$ denote the set of all points ξ_0 in the ξ-plane for which there exists a point u_0 in E such that $\xi(u_0) = \xi_0$. If $\overline{E}$ be any set in the ξ-plane, let $T^{-1}(\overline{E})$ denote the set of all points u_0 in $\mathfrak{D}$ such that $\xi(u_0)$ is in $\overline{E}$. If T_1:$[\xi_1(u), \mathfrak{D}_1]$; T_2:$[\xi_2(u), \mathfrak{D}_2]$ are two bounded continuous transformations, then their distance $\rho(T_1, T_2; E)$ on any set E in both $\mathfrak{D}_1$ and $\mathfrak{D}_2$ is the least upper bound of $\|\xi_1(u) - \xi_2(u)\|$ for u in E. For any set E in the u-plane, and for any point ξ_0 in the ξ-plane, $N(\xi_0, T, E)$ is defined to be the number (possibly $+\infty$) of points in the set $T^{-1}(\xi_0) \cdot E$. For fixed ξ and T, $N(\xi, T, E)$ is a non-negative completely additive set function.

2. If $\mathfrak{R}$ be any Jordan region in $\mathfrak{D}$, and if k be any non-negative integer, define $\overline{\mathfrak{K}}(k, T, \mathfrak{R})$ to be the set of those points ξ_0 in the ξ-plane for which there exists a positive number ϵ such that $N(\xi_0, T_*, \mathfrak{R}) \geqq k$ for every bounded continuous transformation T_* satisfying $\rho(T_*, T; \mathfrak{R}) < \epsilon$. Clearly $\overline{\mathfrak{K}}(k, T, \mathfrak{R})$ contains $\overline{\mathfrak{K}}(k+1, T, \mathfrak{R})$ for $k = 0, 1, 2, \cdots$. Define

$$\overline{\mathfrak{K}}(\infty, T, \mathfrak{R}) = \prod_{k=0}^{\infty} \overline{\mathfrak{K}}(k, T, \mathfrak{R}).$$

A function $\mathrm{K}(\xi, T, \mathfrak{R})$ is defined by the relations

$$\mathrm{K}(\xi, T, \mathfrak{R}) = \begin{cases} k & \text{on } \overline{\mathfrak{K}}(k, T, \mathfrak{R}) - \overline{\mathfrak{K}}(k+1, T, \mathfrak{R}); \\ +\infty & \text{on } \overline{\mathfrak{K}}(\infty, T, \mathfrak{R}). \end{cases}$$

(4) The introduction of the R² paper contains a summary of their results, together with reference to the location of the proofs.

Given a domain $\mathcal{D}$ in $\mathfrak{D}$ let $\mathfrak{R}_n$ be a sequence of Jordan regions whose interiors fill up $\mathcal{D}$ from the interior (see **4**). For fixed ξ and T, the sequence $K(\xi, T, \mathfrak{R}_n)$ has a limit (possibly $+\infty$) which is independent of the choice of the sequence of regions $\mathfrak{R}_n$ whose interiors fill up $\mathcal{D}$. This limit is denoted by $K(\xi, T, \mathcal{D})$, and is termed the *essential multiplicity* of ξ under T with respect to $\mathcal{D}$. It has the following properties.

1. The essential multiplicity $K(\xi, T, \mathcal{D})$, for fixed T and $\mathcal{D}$, is a lower semi-continuous function of ξ.

2. If $\mathcal{D}_n$ is a sequence of domains filling up $\mathcal{D}$ from the interior, if T_n: $[\xi_n(u), \mathcal{D}_n]$ is a sequence of bounded continuous transformations such that for every closed set F in $\mathcal{D}$ it is true that $\lim \rho(T_n, T; F)=0$, then for every ξ one has $\liminf K(\xi, T_n, \mathcal{D}_n) \geqq K(\xi, T, \mathcal{D})$. In particular, if T_n is given by $[\xi(u), \mathcal{D}_n]$ then $\lim K(\xi, T_n, \mathcal{D}_n)=K(\xi, T, \mathcal{D})$.

3. For any Jordan region $\mathfrak{R}$ in $\mathfrak{D}$, it is true that $K(\xi, T, \mathfrak{R}^0)=K(\xi, T, \mathfrak{R})$.

3. A set $\mathcal{B}$ in the u-plane is termed a *base set* for the transformation T if it is measurable, and for every closed oriented square s whose interior s^0 is in $\mathfrak{D}$, the set $T(s^0\cdot\mathcal{B})$ is measurable. Let $\mathcal{B}$ be any base set for T. Define, for any closed oriented square s whose interior s^0 is in $\mathfrak{D}$

$$\mathcal{G}(s, T, \mathcal{B}) = | T(s^0\cdot\mathcal{B}) |. \tag{1}$$

The transformation T is said to be of *bounded variation* with respect to the base set $\mathcal{B}$—briefly, BV $\mathcal{B}$—if there exists a finite positive constant M such that for any finite system of nonoverlapping, closed, oriented squares s_i whose interiors are in $\mathfrak{D}$, it is true that

$$\sum \mathcal{G}(s_i, T, \mathcal{B}) < M.$$

The transformation T is said to be *absolutely continuous* with respect to the base set $\mathcal{B}$—briefly, AC $\mathcal{B}$—if for every positive number ϵ, there exists a positive number η_ϵ such that

$$\sum \mathcal{G}(s_i, T, \mathcal{B}) < \epsilon$$

for every finite system of nonoverlapping, closed, oriented squares s_i whose interiors are in $\mathfrak{D}$ and for which

$$\sum | s_i | < \eta_\epsilon.$$

If T is AC $\mathcal{B}$, then it follows that T is BV $\mathcal{B}$.

LEMMA. *If $\mathcal{B}_1$ and $\mathcal{B}_2$ are base sets for T, then $\mathcal{B}=\mathcal{B}_1+\mathcal{B}_2$ is a base set for T. A necessary and sufficient condition that T be* BV $\mathcal{B}$ *is that T be both* BV $\mathcal{B}_1$ *and* BV $\mathcal{B}_2$. *A necessary and sufficient condition that T be* AC $\mathcal{B}$ *is that T be both* AC $\mathcal{B}_1$ *and* AC $\mathcal{B}_2$.

4. A necessary and sufficient condition that a bounded continuous trans-

formation T be BV $\mathcal{B}$ is that $N(\xi, T, \mathfrak{D}\cdot\mathcal{B})$ be summable[8]. If T is BV $\mathcal{B}$, then $N(\xi, T, E\cdot\mathcal{B})$ is measurable and summable for any open or closed set E relative to $\mathfrak{D}$. If T is BV $\mathcal{B}$, then the function of squares $\mathcal{G}(s, T, \mathcal{B})$ defined in II, 3, 1 possesses a derivative $D(u, T, \mathcal{B})$ almost everywhere in $\mathfrak{D}$. This derivative is summable in $\mathfrak{D}$, and one has, on every open set O in $\mathfrak{D}$[9]

$$\int_O D(u, T, \mathcal{B})du \leqq \int N(\xi, T, O\cdot\mathcal{B})d\xi.$$

If T is AC $\mathcal{B}$, then the sign of equality holds here. Conversely, if the sign of equality holds here for $O\equiv\mathfrak{D}$, then it holds for every open set O in $\mathfrak{D}$, and T is AC $\mathcal{B}$. If T is AC $\mathcal{B}$, and if E is any measurable set in $\mathfrak{D}$, then $N(\xi, T, E\cdot\mathcal{B})$ is measurable and summable; thus if $H(\xi)$ is a finite-valued, measurable function, then $H(\xi)N(\xi, T, E\cdot\mathcal{B})$ is a measurable function. Under these conditions, it is also true that $H(\xi(u))\ D(u, T, \mathcal{B})$ is measurable in $\mathfrak{D}$. Finally, if the transformation T is AC $\mathcal{B}$; the function $H(\xi)$ is finite-valued and measurable; the set E in $\mathfrak{D}$ is measurable; one of $H(\xi)N(\xi, T, E, \mathcal{B})$, $H(\xi(u))D(u, T, \mathcal{B})$ is summable, then both of these functions are summable, and

$$\text{1.} \qquad \int_E H(\xi(u))D(u, T, \mathcal{B})du = \int H(\xi)N(\xi, T, E\cdot\mathcal{B})d\xi.$$

5. Rado and Reichelderfer studied closely the notions of BV $\mathcal{B}$ and AC $\mathcal{B}$ for a certain choice of the base set $\mathcal{B}$ which is now described. Let $\mathfrak{R}$ be any Jordan region in $\mathfrak{D}$ (see **4**). Let ξ_0 be any point in the ξ-plane not on the image under T of the boundary of $\mathfrak{R}$. If C be one of the curves bounding $\mathfrak{R}$, then the image of C under T, taken as u traverses C in a positive sense relative to $\mathfrak{R}$, is a directed, closed continuous curve $\overline{C}$ not passing through ξ_0; consequently ξ_0 has a well defined topological index with respect to $\overline{C}$. The sum of these indices taken over all the boundary curves of $\mathfrak{R}$ is denoted by $\mu(\xi_0, T, \mathfrak{R})$. For points ξ_* on the image under T of the boundary of $\mathfrak{R}$, one puts $\mu(\xi_*, T, \mathfrak{R})=0$.

6. Let ξ_0 be any point in the ξ-plane. The set $T^{-1}(\xi_0)$ is a closed set relative to $\mathfrak{D}$, hence decomposes in a unique way into components which are maximal connected closed sets relative to $\mathfrak{D}$. If a component of $T^{-1}(\xi_0)$ has a positive distance from the boundary of $\mathfrak{D}$, then it is a connected closed set in the absolute sense—that is, a continuum; such a component is termed a maximal model continuum for ξ_0 under T in $\mathfrak{D}$, and is denoted generically by $\sigma(\xi_0, T)$. A $\sigma(\xi_0, T)$ is termed essential if in every open neighborhood of $\sigma(\xi_0, T)$, there is a Jordan region $\mathfrak{R}$ containing $\sigma(\xi_0, T)$ in its interior and for

(8) Since all functions considered in the ξ-plane are zero outside a sufficiently large disc, they are termed summable whenever they are summable on such a disc, and no range of integration will be explicitly indicated.

(9) See Footnote 5.

which $\mu(\xi_0, T, \mathfrak{R})$ is not zero. If $\mathcal{D}$ be any subdomain in $\mathfrak{D}$, then the number of essential maximal model continua $\sigma(\xi_0, T)$ for ξ_0 under T in $\mathcal{D}$ is equal to the essential multiplicity $\mathrm{K}(\xi_0, T, \mathcal{D})$ (see II, 2). If $\mathfrak{R}$ be any Jordan region in $\mathfrak{D}$ for which $\mu(\xi_0, T, \mathfrak{R})$ is not zero, then ξ_0 has an essential maximal model continuum under T in the interior of $\mathfrak{R}$.

7. If $\mathcal{D}$ be any domain in $\mathfrak{D}$, then denote by $E(T, \mathcal{D})$ the set of all points u_0 which belongs to some essential maximal model continuum for $\xi(u_0)$ under T in $\mathcal{D}$. Denote by $\mathcal{E}(T, \mathcal{D})$ that subset of $E(T, \mathcal{D})$ which consists of all those points u_0 which themselves constitute essential maximal model continua for $\xi(u_0)$ under T in $\mathcal{D}$; evidently $\mathcal{E}(T, \mathcal{D}) = \mathcal{E}(T, \mathfrak{D}) \cdot \mathcal{D}$, but a similar formula does *not* generally hold for $E(T, \mathcal{D})$. If u_0 be any point of $\mathcal{E}(T, \mathfrak{D})$ which has a neighborhood free of points belonging to other essential maximal model continua for $\xi(u_0)$ under T in $\mathfrak{D}$, then it is true that $\mu(\xi(u_0), T, \mathfrak{R})$ has a nonzero value independent of the choice of a Jordan region $\mathfrak{R}$ in this neighborhood which contains u_0 in its interior and whose boundary contains no point of $T^{-1}(\xi(u_0))$; denote this value by $j(u_0, T)$. For all points u_* in $\mathfrak{D}$ not having the properties of u_0, set $j(u_*, T)$ equal to zero. Then $j(u, T)$ is a Baire function in $\mathfrak{D}$ and $|j(u, T)|$ does not exceed one, except possibly on a denumerable set of points in $\mathfrak{D}$.

8. It is the set $\mathcal{E}(T, \mathfrak{D})$ which Rado and Reichelderfer employ for a base set (see II, 3); it is the set $E(T, \mathfrak{D})$ which plays a prominent role in the following theory. Because the results for these two base sets are so closely related, the results for the set $\mathcal{E}(T, \mathfrak{D})$ developed by Rado and Reichelderfer are now summarized as a basis for stating and proving results for the set $E(T, \mathfrak{D})$. Let K_0 denote the class of all bounded continuous transformations $T: [\xi(u), \mathfrak{D}]$ which are BV $\mathcal{E}(T, \mathfrak{D})$. Let K_1 denote that subclass of K_0 (see II, 3) consisting of all transformations T which are AC $\mathcal{E}(T, \mathfrak{D})$. Let K_2 denote the class of all transformations T in K_1 for which the relation $N(\xi, T, \mathcal{E}(T, \mathfrak{D})) = \mathrm{K}(\xi, T, \mathfrak{D})$ holds almost everywhere. Finally, let K_3 denote the class of all transformations T in K_2 for which the ordinary Jacobian

$$J(u, T) = \partial(\xi^1, \xi^2)/\partial(u^1, u^2), \qquad u = (u^1, u^2) \text{ in } \mathfrak{D},$$

exists almost everywhere in $\mathfrak{D}$. If T is in K_0, then $D(u, T, \mathcal{E}(T, \mathfrak{D}))$ exists almost everywhere in $\mathfrak{D}$ and is summable on $\mathfrak{D}$ (see II, 3)—denote this derivative by $D(u, T)$. Define

$$\mathcal{J}(u, T) = \begin{cases} j(u, T)D(u, T) & \text{wherever } D(u, T) \text{ exists;} \\ 0 & \text{otherwise.} \end{cases}$$

The function $\mathcal{J}(u, T)$ is termed the ***generalized Jacobian*** for the transformation T. From II, 7 it follows that $|\mathcal{J}(u, T)| \leq D(u, T)$ almost everywhere in $\mathfrak{D}$, and $\mathcal{J}(u, T)$ is measurable in $\mathfrak{D}$; hence $\mathcal{J}(u, T)$ is summable in $\mathfrak{D}$. If T is in the class K_2, it follows that $|\mathcal{J}(u, T)| = D(u, T)$ almost everywhere in $\mathfrak{D}$. If T is in the class K_3, it is true that $\mathcal{J}(u, T) = J(u, T)$ almost everywhere in $\mathfrak{D}$.

The class K_3 contains all bounded continuous transformations $T:[\xi(u), \mathfrak{D}]$ which satisfy a Lipschitz condition in the following restricted sense: there exists a finite constant L such that if u_1 and u_2 are any two points of $\mathfrak{D}$ for which the line segment joining them is contained in $\mathfrak{D}$, then $\|\xi(u_1)-\xi(u_2)\| \leqq L\|u_1-u_2\|$. If T is a bounded continuous transformation for which the ordinary Jacobian exists almost everywhere in $\mathfrak{D}$, and for which $\mathrm{K}(\xi, T, \mathfrak{D})$ is summable, then $D(u, T) \geqq |J(u, T)|$ almost everywhere in $\mathfrak{D}$(7).

9. Combining results stated in II, **4**, **8**, one obtains the

1. Lemma. *Let $T:[\xi(u), \mathfrak{D}]$ be a bounded continuous transformation for which $\mathrm{K}(\xi, T, \mathfrak{D})$ is summable. Then $\mathcal{J}(u, T)$ exists almost everywhere in $\mathfrak{D}$, is summable on $\mathfrak{D}$, and*

$$\int \mathrm{K}(\xi, T, \mathcal{D})d\xi \geqq \int N(\xi, T, \mathcal{E}(T, \mathcal{D}))d\xi \geqq \int_{\mathcal{D}} D(u, T)du \geqq \int_{\mathcal{D}} |\mathcal{J}(u, T)|\, du$$

for every domain $\mathcal{D}$ in $\mathfrak{D}$. A necessary and sufficient condition that all the signs of equality hold here for $\mathcal{D}\equiv\mathfrak{D}$ is that T be in the class K_2. When T is in K_2, all the signs of equality also hold for every domain $\mathcal{D}$ in $\mathfrak{D}$.

2. Lemma. *Let $T:[\xi(u), \mathfrak{D}]$ be a bounded continuous transformation for which $\mathrm{K}(\xi, T, \mathfrak{D})$ is summable, and for which the ordinary Jacobian $J(u, T)$ exists almost everywhere in $\mathfrak{D}$. Then $J(u, T)$ is summable on $\mathfrak{D}$, and*

$$\int \mathrm{K}(\xi, T, \mathcal{D})d\xi \geqq \int N(\xi, T, \mathcal{E}(T, \mathcal{D}))d\xi \geqq \int_{\mathcal{D}} D(u, T)du \geqq \int_{\mathcal{D}} |J(u, T)|\, du$$

for every domain $\mathcal{D}$ in $\mathfrak{D}$. A necessary and sufficient condition that all the signs of equality hold here for $\mathcal{D}\equiv\mathfrak{D}$ is that T be in the class K_3, When T is in K_3, all the signs of equality also hold for every domain $\mathcal{D}$ in $\mathfrak{D}$.

10. Lemma. *If $T:[\xi(u), \mathfrak{D}]$ be any bounded continuous transformation, then the set $E(T, \mathfrak{D})$ (see II, 7) is a product of open sets, hence a Borel set.*

Proof. Let n be any positive integer. Denote by E_n the set of points u_0 in $\mathfrak{D}$ for each of which there exists a Jordan region $\mathfrak{R}$ in $\mathfrak{D}$ satisfying the following conditions: u_0 lies in the interior of $\mathfrak{R}$; $T(\mathfrak{R})$ lies in the open disc $\|\xi-\xi(u_0)\|<n^{-1}$; $\mu(\xi(u_0), T, \mathfrak{R})\neq 0$. Evidently each E_n is an open set. One easily verifies that $\prod E_n = E(T, \mathfrak{D})$, so the lemma is established.

Set $e(T, \mathfrak{D}) = E(T, \mathfrak{D}) - \mathcal{E}(T, \mathfrak{D})$—that is, $e(T, \mathfrak{D})$ is the set of those points u_0 belonging to some nondegenerate essential maximal model continuum for $\xi(u_0)$ under T in $\mathfrak{D}$. Rado and Reichelderfer have shown that $\mathcal{E}(T, \mathfrak{D})$ is also a product of open sets, hence a Borel set. Thus $e(T, \mathfrak{D})$ is a Borel set. By a general theorem (see Kuratowski [1, p. 249]) it follows that for every choice of an open square s^0 in the u-plane, the sets $T(s^0\cdot e(T, \mathfrak{D}))$, $T(s^0\cdot E(T, \mathfrak{D}))$ are

(7) The last result is established in §5.6 of the R^2 paper.

both measurable. Thus both the sets $e(T, \mathfrak{D})$, $E(T, \mathfrak{D})$ may serve as base sets (see II, 3), and the general theory in II, 3, 4 is applicable.

11. Lemma. *A necessary and sufficient condition that the set $T(e(T, \mathfrak{D}))$ be of measure zero is that T be* BV $e(T, \mathfrak{D})$. *Whenever T is* BV $e(T, \mathfrak{D})$), *it is also* AC $e(T, \mathfrak{D})$.

Proof. Observe that (see II, 1)

$$N(\xi, T, e(T, \mathfrak{D})) = \begin{cases} +\infty & \text{for } \xi \text{ in } T(e(T, \mathfrak{D})); \\ 0 & \text{otherwise.} \end{cases}$$

Thus a necessary and sufficient condition that $N(\xi, T, e(T, \mathfrak{D}))$ be summable is that $|T(e(T, \mathfrak{D}))| = 0$; in view of the facts in II, 4, this establishes the first part of the lemma. If T is BV $e(T, \mathfrak{D})$), it follows at once that

$$0 \leqq \int_{\overline{\mathfrak{D}}} D(u, T, e(T, \mathfrak{D}))du \leqq \int N(\xi, T, e(T, \mathfrak{D}))d\xi = 0,$$

so that the sign of equality holds here, and T is AC $e(T, \mathfrak{D})$.

From this lemma and the lemma in II, 3 follows the

Corollary. *Let T be any bounded continuous transformation which is* BV $e(T, \mathfrak{D})$. *A necessary and sufficient condition that T be* BV $E(T, \mathfrak{D})$ *is that T be* BV $\mathcal{E}(T, \mathfrak{D})$. *A necessary and sufficient condition that T be* AC $E(T, \mathfrak{D})$ *is that T be* AC $\mathcal{E}(T, \mathfrak{D})$.

If T is BV $e(T, \mathfrak{D})$ it is evident that (see II, 3, 7, 10) for every closed oriented square s whose interior is in $\mathfrak{D}$,

$$\mathcal{G}(s, T, E(T, \mathfrak{D})) = \mathcal{G}(s, T, \mathcal{E}(T, \mathfrak{D})).$$

Thus whenever T is BV $e(T, \mathfrak{D})$, it follows that $D(u, T, E(T, \mathfrak{D})) = D(u, T, \mathcal{E}(T, \mathfrak{D})) \equiv D(u, T)$ almost everywhere in $\mathfrak{D}$ (see II, 4, 8).

12. Let $T: [\xi(u), \mathfrak{D}]$ be a bounded continuous transformation. From the definitions in II, 1, 2, 6, 7, it is clear that

1. $$N(\xi, T, E(T, \mathfrak{D})) \geqq \mathrm{K}(\xi, T, \mathfrak{D}) \geqq N(\xi, T, \mathcal{E}(T, \mathfrak{D})),$$

and the sign of inequality holds between any two of these three functions if and only if ξ lies in the set $T(e(T, \mathfrak{D}))$ and one of the functions involved is finite. Since $\mathrm{K}(\xi, T, \mathfrak{D})$ is measurable (see II, 2, 1), it follows that a necessary condition that $\mathrm{K}(\xi, T, \mathfrak{D})$ be summable is that T be BV $\mathcal{E}(T, \mathfrak{D})$, while a sufficient condition that $\mathrm{K}(\xi, T, \mathfrak{D})$ be summable is that T be BV $E(T, \mathfrak{D})$; in this latter case, the signs of equality in relation 1 hold almost everywhere. Now suppose that T is BV $\mathcal{E}(T, \mathfrak{D})$; then since $N(\xi, T, \mathcal{E}(T, \mathfrak{D}))$ is finite almost everywhere, one concludes that a necessary and sufficient condition that $\mathrm{K}(\xi, T, \mathfrak{D}) = N(\xi, T, \mathcal{E}(T, \mathfrak{D}))$ almost everywhere is that T be BV $E(T, \mathfrak{D})$.

In view of the definition of the class K_2 (see II, 8), the corollary in the preceding section, and these remarks, one obtains the

LEMMA. *A necessary and sufficient condition that a bounded continuous transformation* $T:[\xi(u), \mathfrak{D}]$ *be in the class* K_2 *is that* T *be* AC $E(T, \mathfrak{D})$.

13. The lemma in the preceding section gives a simple characterization of transformations in the class K_2 and makes available all the results established by Rado and Reichelderfer for this class whenever T is AC $E(T, \mathfrak{D})$. For example, they have the

CLOSURE THEOREM. *Let there be given bounded domains* $\mathcal{D}$ *and* $\mathcal{D}_n$ *in the u-plane and bounded continuous transformations* $T: [\xi(u), \mathcal{D}]$ *and* $T_n: [\xi_n(u), \mathcal{D}_n]$, $n=1, 2, \cdots$, *with the following properties*: (i) *the domains* $\mathcal{D}_n$ *fill up* $\mathcal{D}$ *from the interior* (*see* 4); (ii) *the generalized Jacobian* $\mathcal{J}(u, T)$ *exists almost everywhere in* $\mathcal{D}$ *and is summable on* $\mathcal{D}$ (*see* II, 8); (iii) T_n *is in* K_2 *for* $n=1, 2, \cdots$; (iv) *for every closed set* F *contained in* $\mathcal{D}$, *it is true that*

$$\lim \rho(T_n, T; F) = 0, \qquad \lim \int_F | \mathcal{J}(u, T_n) - \mathcal{J}(u, T) | \, du = 0.$$

Then T *is in* K_2.

Using the preceding results, this theorem may be restated and improved as follows.

14. MODIFIED CLOSURE THEOREM. *Let there be given bounded domains* $\mathcal{D}$ *and* $\mathcal{D}_n$ *in the u-plane and bounded continuous transformations* $T: [\xi(u), \mathcal{D}]$ *and* $T_n: [\xi_n(u), \mathcal{D}_n]$, $n=1, 2, \cdots$, *with the following properties*: (i) *the domains* $\mathcal{D}_n$ *fill up* $\mathcal{D}$ *from the interior*; (ii) *the generalized Jacobian* $\mathcal{J}(u, T)$ *exists almost everywhere in* $\mathcal{D}$ *and is summable on* $\mathcal{D}$; (iii) T_n *is* AC $E(T_n, \mathcal{D}_n)$ *for* $n=1, 2, \cdots$; (iv′) *there exists a sequence of Jordan regions* $\mathfrak{R}_m$ *in* $\mathcal{D}$ *whose interiors* $\mathfrak{R}_m^0$ *fill up* $\mathcal{D}$ *from the interior, and such that*

$$\lim_{n\to\infty} \rho(T_n, T; \mathfrak{R}_m) = 0, \qquad \lim_{n\to\infty} \int_{\mathfrak{R}_m} | \mathcal{J}(u, T_n) | \, du = \int_{\mathfrak{R}_m} | \mathcal{J}(u, T) | \, du, \qquad m = 1, 2, \cdots .$$

Then T *is* AC $E(T, \mathcal{D})$.

Inspection of the proof of Rado and Reichelderfer for their closure theorem reveals that property (iv′), which is a consequence of (iv), is all that is needed for that proof.

COROLLARY. *Condition* (iv′) *of the above theorem may be replaced by the following condition*: (iv″) *for every closed oriented square* s *contained in* $\mathcal{D}$, *it is true that*

$$\lim \rho(T_n, T; s) = 0, \qquad \lim \int_s |\mathcal{J}(u, T_n)|\, du = \int_s |\mathcal{J}(u, T)|\, du.$$

Proof. Let R denote any Jordan region in $\mathcal{D}$ which may be expressed as a sum of a finite number of nonoverlapping closed oriented squares s. From (iv″) it follows that

$$\lim \rho(T_n, T; R) = 0, \qquad \lim \int_R |\mathcal{J}(u, T_n)|\, du = \int_R |\mathcal{J}(u, T)|\, du.$$

By a known theorem in topology (see Kerékjártó [1]), there exists a sequence of regions R_n whose interiors fill up $\mathcal{D}$ from the interior. Thus condition (iv″) implies (iv′), and the corollary is established.

15. In view of the lemma in II, **12**, the class K_3 defined in II, **8** may be characterized as the class of all transformations $T: [\xi(u), \mathfrak{D}]$ which are AC $E(T, \mathfrak{D})$ and for which the ordinary Jacobian $J(u, T)$ exists almost everywhere in $\mathfrak{D}$. For the class K_3 another closure theorem is given by Rado and Reichelderfer, the proof of which is based upon the closure theorem for the class K_2 stated in II, **13**. By using the modified closure theorem just given, one may parallel their proof to establish the

Modified closure theorem. *Let there be given bounded domains $\mathcal{D}$ and $\mathcal{D}_n$ and bounded continuous transformations $T: [\xi(u), \mathcal{D}]$ and $T_n: [\xi_n(u), \mathcal{D}_n]$, $n=1, 2, \cdots$, with the following properties*: (i) *the domains $\mathcal{D}_n$ fill up $\mathcal{D}$ from the interior*; (ii) *the ordinary Jacobian $J(u, T)$ exists almost everywhere in $\mathcal{D}$ and is summable on $\mathcal{D}$*; (iii) *T_n is* AC $E(T_n, \mathcal{D}_n)$ *and the ordinary Jacobians $J(u, T_n)$ exist almost everywhere in $\mathcal{D}_n$ for $n=1, 2, \cdots$*; (iv′) *there exists a sequence of Jordan regions $\mathfrak{R}_m$ in $\mathcal{D}$ whose interiors $\mathfrak{R}_m^0$ fill up $\mathcal{D}$ from the interior, and such that*

$$\lim_{n\to\infty} \rho(T_n, T; \mathfrak{R}_m) = 0, \qquad \lim_{n\to\infty} \int_{\mathfrak{R}_m} |J(u, T_n)|\, du = \int_{\mathfrak{R}_m} |J(u, T)|\, du,$$
$$m = 1, 2, \cdots.$$

Then T is AC $E(T, \mathcal{D})$.

Corollary. *Condition* (iv′) *of the above theorem may be replaced by the following condition*: (iv″) *for every closed oriented square s contained in $\mathcal{D}$, it is true that*

$$\lim \rho(T_n, T; s) = 0, \qquad \lim \int_s |J(u, T_n)|\, du = \int_s |J(u, T)|\, du.$$

If in these results, condition (ii) be weakened by dropping the requirement that the ordinary Jacobians $J(u, T_n)$ exist almost everywhere in $\mathcal{D}_n$ for $n=1, 2, \cdots$, and if these Jacobians be replaced by the generalized Jacobians $\mathcal{J}(u, T_n)$ for $n=1, 2, \cdots$, then the conclusions remain the same.

16. The results to be established in the following sections are necessary for a comparison of certain notions in this paper with those in the literature. In the theory of bounded continuous transformations just sketched, the range of definition has been a bounded domain $\mathfrak{D}$ (see I, 1). In applications, one may have a continuous transformation given by $[\xi(u), \mathfrak{B}]$, where $\mathfrak{B}$ is a simple Jordan region (see 4) and $\xi(u)$ is a pair of functions defined and continuous on the closed set $\mathfrak{B}$. Evidently the transformation given by $[\xi(u), \mathfrak{B}^0]$ is then bounded and continuous. From work of Rado and Reichelderfer (see II, 2, 6), it follows that the transformations $[\xi(u), \mathfrak{B}]$ and $[\xi(u), \mathfrak{B}^0]$ have the same essential multiplicity functions and the same essential maximal model continua. Since the essential multiplicity and the essential maximal model continua play the basic role in this paper—that is, since the transformations behave essentially alike—there will be no confusion if one designates either of them by T. In the sequel this is done, but it is to be understood that whenever preceding theory is applied, T is to be interpreted as the transformation $[\xi(u), \mathfrak{B}^0]$.

17. Lemma. *Let T: $[\xi(u), \mathfrak{D}]$ be a bounded continuous transformation which is* BV $\mathcal{B}$, *where $\mathcal{B}$ is an arbitrary base set (see* II, 3). *Given a positive number ϵ, it is true that the number of mutually exclusive sets E, each of which is either an open set or a closed set relative to $\mathfrak{D}$, for which the measure of $T(E\cdot\mathcal{B})$ exceeds ϵ is finite.*

Proof. Let $E_1, \cdots, E_m$ be any finite system of mutually exclusive sets having the required properties. Since T is BV $\mathcal{B}$, it follows that $N(\xi, T, \mathfrak{D}\cdot\mathcal{B})$, $N(\xi, T, E_i\cdot\mathcal{B})$ are measurable and summable (see II, 4). The lemma follows from the inequalities

$$\epsilon^{-1}\int N(\xi, T, \mathfrak{D}\cdot\mathcal{B})d\xi \geqq \epsilon^{-1}\sum_{i=1}^{m}\int N(\xi, T, E_i\cdot\mathcal{B})d\xi \geqq \epsilon^{-1}\sum_{i=1}^{m}\left| T(E_i\cdot\mathcal{B})\right| > m.$$

From this lemma comes the

Corollary. *Let T:$[\xi(u), \mathfrak{D}]$ be a bounded continuous transformation which is* BV $\mathcal{B}$. *Then the number of lines l in any family of parallel lines in the u-plane for which $T(l\cdot\mathcal{B})$ has positive measure is at most enumerable.*

18. Let T: $[\xi(u), \mathfrak{B}]$ where $\mathfrak{B}$ is a simple Jordan region in the u-plane, be a continuous transformation. Let $\overline{\mathfrak{B}}$ be any other simple Jordan region in the u-plane, and consider any topological map of $\overline{\mathfrak{B}}$ onto $\mathfrak{B}$ given by $[\bar{u}(u), \overline{\mathfrak{B}}]$. Denote by $\overline{T}$ the continuous transformation $[\bar{\xi}(u), \overline{\mathfrak{B}}]$ where $\bar{\xi}(u)=\xi(\bar{u}(u))$ for u in $\overline{\mathfrak{B}}$. Since the sets $E(T, \mathfrak{B}^0)$ and $E(\overline{T}, \overline{\mathfrak{B}}^0)$ are in biunique correspondence under the map, it is true that $N(\xi, T, E(T, \mathfrak{B}^0))=N(\xi, \overline{T}, E(\overline{T}, \overline{\mathfrak{B}}^0))$. This implies (see II, 4) the

Lemma. *If T is* BV $E(T, \mathfrak{B}^0)$, *then $\overline{T}$ is* BV $E(\overline{T}, \overline{\mathfrak{B}}^0)$.

19. Now let $T_0[\xi(u), \mathfrak{B}]$, where $\mathfrak{B}$ is a simple Jordan region in the u-plane, be a continuous transformation for which $K(\xi, T, \mathfrak{B}^0)$ is summable (see II, **2**). If $S(\mathfrak{B})$ be any finite system (see **4**), denote by C the set of points in $\mathfrak{B}^0$ but not in the interior of any B belonging to $S(\mathfrak{B})$. Then (see II, **6**), wherever $K(\xi, T, \mathfrak{B}^0)$ is finite—hence almost everywhere,

$$1. \quad K(\xi, T, \mathfrak{B}^0)\begin{cases} = \sum K(\xi, T, B^0) \text{ for } B \text{ in } S(\mathfrak{B}) \text{ if } \xi \text{ is not in } T(C\cdot E(T, \mathfrak{B}^0)); \\ > \sum K(\xi, T, B^0) \text{ for } B \text{ in } S(\mathfrak{B}) \text{ if } \xi \text{ is in } T(C\cdot E(T, \mathfrak{B}^0)). \end{cases}$$

Thus

$$2. \qquad \int K(\xi, T, \mathfrak{B}^0)d\xi \geqq \sum \int K(\xi, T, B^0) \quad \text{for} \quad B \text{ in } S(\mathfrak{B}),$$

and a necessary and sufficient condition that the sign of equality hold here is that the sign of equality hold almost everywhere in relation 1 —that is, that the set $T(C\cdot E(T, \mathfrak{B}^0))$ be of measure zero. For brevity, any finite system $S(\mathfrak{B})$ for which the sign of equality holds in relation 2 is termed a maximal system for T.

20. Lemma. *Given a continuous transformation T: $[\xi(u), \mathfrak{B}]$. A necessary and sufficient condition that, for any positive number δ there exist maximal systems $S(\mathfrak{B})$ for T such that $\|S(\mathfrak{B})\|$ is less than δ is that T be* BV $E(T, \mathfrak{B}^0)$.

Proof. First, assume that there exists a sequence of finite systems $S_n(\mathfrak{B})$ such that $\|S_n(\mathfrak{B})\|$ converges to zero and

$$\int K(\xi, T, \mathfrak{B}^0)d\xi = \sum \int K(\xi, T, B^0)d\xi \quad \text{for} \quad B \text{ in } S_n(\mathfrak{B}), \qquad n = 1, 2, \cdots.$$

Denote C_n the set of points in $\mathfrak{B}^0$ none of which is in the interior of a region B belonging to $S_n(\mathfrak{B})$ for $n=1, 2, \cdots$; set $\Gamma=\sum C_n$. From the remark in II, **19**, it follows that the set $T(\Gamma\cdot E(T, \mathfrak{B}^0))$ is of measure zero. On the other hand, the set $T(e(T, \mathfrak{B}^0))$ (see II, **10**) is clearly a subset of $T(\Gamma\cdot E(T, \mathfrak{B}^0))$, hence also of measure zero. By the lemma in II, **11**, it is true that T is BV $e(T, \mathfrak{B}^0)$. Since $K(\xi, T, \mathfrak{B}^0)$ is summable, it follows from II, **12** that T is BV $\mathcal{E}(T, \mathfrak{B}^0)$. From the lemma in II, **3**, it is clear that T is BV $E(T, \mathfrak{B}^0)$. This establishes the necessity of the condition. Next, assume that T is BV $E(T, \mathfrak{B}^0)$; then $K(\xi, T, \mathfrak{B}^0)$ is summable (see II, **12**). First, consider the special case when the simple Jordan region $\mathfrak{B}$ is an interval $\mathfrak{J}$ (see **4**). Given a positive number δ, there exists, according to the corollary in II, **17**, an interval subdivision $S(\mathfrak{J})$ such that $\|S(\mathfrak{J})\|$ is less than δ and $T(l\cdot E(T, \mathfrak{J}^0))$ is of measure zero for every line l forming the subdivision. Denote by C the set of points in $\mathfrak{J}^0$ belonging to the lines of subdivision forming $S(\mathfrak{J})$. Clearly $T(C\cdot E(T, \mathfrak{J}^0))$ is of measure zero, so that $S(\mathfrak{J})$ is a maximal system for T with $\|S(\mathfrak{J})\|$ less than δ. Now consider the general case when

$\mathfrak{B}$ is any simple Jordan region. Let $\mathfrak{J}$ be any interval in the u-plane; then there exists a topological map of $\mathfrak{J}$ onto $\mathfrak{B}$ given by $[\bar{u}(u), \mathfrak{J}]$. Denote by $\overline{T}$ the transformation $[\bar{\xi}(u), \mathfrak{J}]$ where $\bar{\xi}(u) = \xi(\bar{u}(u))$ for u in $\mathfrak{J}$. Since $\overline{T}$ is BV $E(\overline{T}, \mathfrak{J}^0)$ (see II, **18**), it follows by the result just established that there exist interval subdivisions $\overline{S}(\mathfrak{J})$ which are maximal with respect to $\overline{T}$ for which $\|\overline{S}(\mathfrak{J})\|$ is arbitrarily small. A maximal system $\overline{S}(\mathfrak{J})$ for $\overline{T}$ corresponds under the map to a system $S(\mathfrak{B})$ which is easily seen to be maximal for T. In view of the uniform continuity of the topological map, $\|S(\mathfrak{B})\|$ will be less than δ provided $\|\overline{S}(\mathfrak{J})\|$ is chosen sufficiently small. This establishes the lemma.

21. The methods of proof for the preceding lemma yield the

LEMMA. *If T_n: $[\xi_n(u), \mathfrak{B}]$, $n=1, 2, \cdots$, be a sequence of continuous transformations each defined on the simple Jordan region $\mathfrak{B}$, and each* BV $E(T_n, \mathfrak{B})$, *then for every positive number δ there exists a subdivision $S(\mathfrak{B})$ having $\|S(\mathfrak{B})\|$ less than δ, such that $S(\mathfrak{B})$ is maximal for each of the T_n for $n=1, 2, \cdots$. If $\mathfrak{B}$ is an interval then $S(\mathfrak{B})$ may be chosen to be an interval subdivision.*

22. Again, let T: $[\xi(u), \mathfrak{B}]$ be a continuous transformation. For any simple Jordan region B in $\mathfrak{B}$, denote by $c(\mathfrak{J}, T, B^0)$ the characteristic function of the set of points ξ where $\mathrm{K}(\xi, T, B^0)$ is positive (see II, **2**). Then $c(\xi, T, B^0)$ is summable, and

$$\int c(\xi, T, B^0)d\xi = |\, T(E(T, B^0))\,| = |\, \mathfrak{K}(1, T, B)\,| \qquad \text{for } B \text{ in } \mathfrak{B}.$$

Consider any sequence of finite systems $S_n(\mathfrak{B})$ for which $\|S_n(\mathfrak{B})\|$ converges to zero. Clearly

$$\sum_{B \text{ in } S_n(\mathfrak{B})} [\mathrm{K}(\xi, T, B^0) - c(\xi, T, B^0)] \begin{cases} \leqq \mathrm{K}(\xi, T, \mathfrak{B}^0); \\ \to_{n\to\infty} 0 \quad \text{if} \quad \mathrm{K}(\xi, T, \mathfrak{B}^0) \text{ is finite.} \end{cases}$$

Thus if $\mathrm{K}(\xi, T, \mathfrak{B}^0)$ is summable, it follows by a theorem of Lebesgue that one may integrate this sequence termwise to obtain

$$\lim_{n\to\infty} \sum_{B \text{ in } S_n(\mathfrak{B})} \int [\mathrm{K}(\xi, T, B^0) - c(\xi, T, B^0)]d\xi = 0.$$

Combining this result with that in II, **21**, one concludes the

LEMMA. *Let T: $[\xi(u), \mathfrak{B}]$ be a continuous transformation which is* BV $E(T, \mathfrak{B}^0)$. *Then for any sequence of maximal systems $S_n(\mathfrak{B})$ for T such that $\|S_n(\mathfrak{B})\|$ converges to zero, it is true that*

$$\int \mathrm{K}(\xi, T, \mathfrak{B}^0)d\xi = \lim_{n\to\infty} \sum_{B \text{ in } S_n(\mathfrak{B})} \int c(\xi, T, B^0)d\xi.$$

CHAPTER III

ON FLAT CONTINUOUS SURFACES

1. Suppose that S is a flat continuous surface (see I, **3**) lying in the ξ-plane. Every representation $[\xi(u), \mathfrak{B}]$ for S determines a bounded continuous transformation T from the simple Jordan region $\mathfrak{B}$ to the ξ-plane (see II, **16**). For brevity, $[\xi(u), \mathfrak{B}]$ is termed BV E, AC E, and so on, whenever T is BV $E(T, \mathfrak{B}^0)$, AC $E(T, \mathfrak{B}^0)$, and so on.

LEMMA. *The essential multiplicity* $K(\xi, T, \mathfrak{B}^0)$ *is independent of the choice of the representation* $[\xi(u), \mathfrak{B}]$ *for* S.

Proof. Let $[\xi_1(u), \mathfrak{B}_1]$, $[\xi_2(u), \mathfrak{B}_2]$ be any two representations for S; denote the corresponding transformations by T_1, T_2. From I, **1** it follows that there exists a sequence of topological maps of $\mathfrak{B}_1$ onto $\mathfrak{B}_2$ given by $[\bar{u}_n(u), \mathfrak{B}_1]$, $n=3, 4, \cdots$, such that $\|\xi_1(u)-\xi_2(\bar{u}_n(u))\|<n^{-1}$ for u in $\mathfrak{B}_1$. Let T_n denote the transformation given by $[\xi_2(\bar{u}_n(u)), \mathfrak{B}_1]$ for $n=3, 4, \cdots$. From the definition of the essential multiplicity (see II, **2**, **6**) and the nature of a topological map, it is clear that $K(\xi, T_2, \mathfrak{B}_2^0)=K(\xi, T_n, \mathfrak{B}_1^0)$ for $n=3, 4, \cdots$. Since (see II, **1**) $\lim \rho(T_n, T_1; \mathfrak{B}_1)=0$, it follows from II, **2**, **2** that $\liminf K(\xi, T_n, \mathfrak{B}_1^0) \geqq K(\xi, T_1, \mathfrak{B}_1^0)$. Combining these relations, one obtains $K(\xi, T_2, \mathfrak{B}_2^0) \geqq K(\xi, T_1, \mathfrak{B}_1^0)$. By symmetry, the opposite inequality follows; thus the lemma is established.

2. In view of the lemma above, one may define an *essential multiplicity* for a flat continuous surface S lying in the ξ-plane by the relation $K(\xi, S) = K(\xi, T, \mathfrak{B}^0)$, where T is the transformation associated with any representation $[\xi(u), \mathfrak{B}]$ for S.

LEMMA. *The essential multiplicity* $K(\xi, S)$ *is a lower semi-continuous functional in each of its arguments* ξ *and* S.

Proof. The fact that $K(\xi, S)$ is a lower semi-continuous function of ξ follows from II, **2**, 1. Next, suppose that the flat surfaces S_n in the ξ-plane converge to the surface S_0. From I, **1** it follows that there exist representations $[\xi_n(u), \mathfrak{B}]$ for S_n, for $n=0, 1, 2, \cdots$, such that the corresponding transformations T_n satisfy $\lim \rho(T_n, T_0; \mathfrak{B})=0$. So from II, **2**, **2** follows

$$\liminf K(\xi, S_n) = \liminf K(\xi, T_n, \mathfrak{B}^0) \geqq K(\xi, T_0, \mathfrak{B}^0) = K(\xi, S_0).$$

This establishes the lemma.

3. For any flat surface S in the ξ-plane, define (see II, **2**)

$$eV(S) = \begin{cases} \int K(\xi, S)d\xi & \text{if } K(\xi, S) \text{ is summable;} \\ +\infty & \text{otherwise.} \end{cases}$$

The quantity $eV(S)$ is termed the *essential variation* for the surface S. If

$eV(S)$ is finite, then S is said to be a surface of *bounded essential variation*—briefly, BEV. From the remarks in II, 12 follow the

1. Corollary. *A sufficient condition that a flat surface be* BEV *is that it possess a representation which is* BV E. *A necessary condition that a flat surface be* BEV *is that each of its representations be* BV $\mathcal{E}$. *Thus if one representation for the surface is* BV E, *then all representations are* BV $\mathcal{E}$.

2. Corollary. *The essential variation* $eV(S)$ *is a lower semi-continuous functional of* S.

4. From the lemmas in II, 9, 12 one obtains the

1. Lemma. *If* S *is a flat surface which is* BEV, *and if* $[\xi(u), \mathfrak{B}]$ *is any representation for* S, *then*

$$eV(S) \geqq \int_{\mathfrak{B}^0} |\mathcal{J}(u, T)|\, du,$$

where T *is the corresponding transformation. A necessary and sufficient condition that the sign of equality hold here is that* $[\xi(u), \mathfrak{B}]$ *be* AC E.

2. Lemma. *If* S *is a flat surface which is* BEV, *and if* $[\xi(u), \mathfrak{B}]$ *is any representation for* S *for which the ordinary Jacobian exists almost everywhere in* $\mathfrak{B}^0$, *then*

$$eV(S) \geqq \int_{\mathfrak{B}^0} |J(u, T)|\, du,$$

where T *is the corresponding transformation. A necessary and sufficient condition that the sign of equality hold here is that* $[\xi(u), \mathfrak{B}]$ be AC E.

5. For the purpose of comparing results with those in the literature for surfaces given in non-parametric representation, it is necessary to recall the concepts for bounded variation and absolute continuity used by Tonelli (see Tonelli [1]). Let $f(u) = f(u^1, u^2)$ be a real, single-valued function defined and continuous on the interval $\mathfrak{J} = [\alpha, \beta] = [\alpha^1, \beta^1; \alpha^2, \beta^2]$ (see 4). For fixed u^2 in $[\alpha^2, \beta^2]$, denote by $V_{u^1}(f; u^2)$ the total variation of $f(u^1, u^2)$ as a function of u^1 on $[\alpha^1, \beta^1]$. The function $V_{u^1}(f; u^2)$ is a lower semi-continuous function of u^2 on $[\alpha^2, \beta^2]$, hence is measurable. Define $V_{u^2}(f; u^1)$ by interchanging the roles of u^1 and u^2. If both $V_{u^1}(f; u^2)$, $V_{u^2}(f; u^1)$ are summable on their respective intervals of definition, then $f(u)$ is said to be of bounded variation in the sense of Tonelli on $\mathfrak{J}$—briefly, BV T on $\mathfrak{J}$—and the total u^1- and u^2-variations of $f(u)$ on $\mathfrak{J}$ are defined by

$$V_{u^1}(f) = \int_{\alpha^2}^{\beta^2} V_{u^1}(f; u^2)\, du^2, \qquad V_{u^2}(f) = \int_{\alpha^1}^{\beta^1} V_{u^2}(f; u^1)\, du^1.$$

If $f(u)$ is BV T on $\mathfrak{J}$, then it follows that the partial derivatives $f_{u^1}(u)$, $f_{u^2}(u)$ exist almost everywhere in $\mathfrak{J}$, are summable on $\mathfrak{J}$, and

1. $$V_{u^1}(f) \geqq \int_{\mathfrak{J}} |f_{u^1}(u)|\, du, \qquad V_{u^2}(f) \geqq \int_{\mathfrak{J}} |f_{u^2}(u)|\, du.$$

6. Assume that $f(u)$ is BV T on $\mathfrak{J}$. If moreover, for almost every u^2 in $[\alpha^2, \beta^2]$ it is true that $f(u^1, u^2)$ is an absolutely continuous function of u^1 on $[\alpha^1, \beta^1]$, and if a similar relation holds with the roles of u^1 and u^2 reversed, then $f(u)$ is said to be absolutely continuous in the sense of Tonelli on $\mathfrak{J}$—briefly, AC T on $\mathfrak{J}$. A necessary and sufficient condition that a function $f(u)$ which is BV T on $\mathfrak{J}$ be AC T on $\mathfrak{J}$ is that the sign of equality hold in both of the relations III, **5**, 1.

7. Let $f(u)$ be a real, single-valued function defined and continuous on the interval $\mathfrak{J} = [\alpha, \beta]$. Consider the continuous transformation defined by

$$^1T\colon \quad \xi^1 = u^2, \qquad \xi^2 = f(u^1, u^2), \qquad (u^1, u^2) \text{ in } \mathfrak{J}.$$

Notice that for a fixed $u^2 = \gamma^2$ in $[\alpha^2, \beta^2]$, ^1T gives a linear transformation from the closed linear interval $u^2 = \gamma^2$, $\alpha^1 \leqq u^1 \leqq \beta^1$ to a bounded portion of the line $\xi^1 = \gamma^2$. For this transformation it is known (see Rado [4]) that a necessary and sufficient condition that $f(u^1, \gamma^2)$ be of bounded variation on $[\alpha^1, \beta^1]$ is that $N((\gamma^2, \xi^2), {}^1T, \mathfrak{J}^0)$ be summable as a function of ξ^2 (see **4**; II, 1), and that if $f(u, \gamma^2)$ is of bounded variation on $[\alpha^1, \beta^1]$ then (see III, **5**)

1. $$\int N((\gamma^2, \xi^2), {}^1T, \mathfrak{J}^0)\, d\xi^2 = V_{u^1}(f; \gamma^2), \qquad \alpha^2 \leqq \gamma^2 \leqq \beta^2.$$

Thus it follows that if $V_{u^1}(f; u^2)$ is summable on $[\alpha^2, \beta^2]$, then $\int N((\xi^1, \xi^2), {}^1T, \mathfrak{J}^0)\, d\xi^2$ is a summable function of ξ^1 on $[\alpha^2, \beta^2]$. Now (see R² [1], chap. III]) $N(\xi, {}^1T, \mathfrak{J}^0)$ is measurable in the ξ-plane. So if $V_{u^1}(f; u^2)$ is summable on $[\alpha^2, \beta^2]$, it follows from the theorem of Fubini that $N(\xi, {}^1T, \mathfrak{J}^0)$ is summable, hence (see II, **4**) ^1T is BV $\mathfrak{J}^0$, and (see III, **5**)

2. $$\int N(\xi, {}^1T, \mathfrak{J}^0)\, d\xi = \int d\xi^1 \int N((\xi^1, \xi^2), {}^1T, \mathfrak{J}^0)\, d\xi^2 = \int_{\alpha^2}^{\beta^2} V_{u^1}(f; u^2)\, du^2 = V_{u^1}(f).$$

Conversely, suppose that ^1T is BV $\mathfrak{J}^0$; then (see II, **4**) it is true that $N(\xi, {}^1T, \mathfrak{J}^0)$ is summable, so that by the theorem of Fubini, the function $N((\gamma^2, \xi^2), {}^1T, \mathfrak{J}^0)$ is summable for almost every choice of $\xi^1 = \gamma^2$ in the interval $[\alpha^2, \beta^2]$, and

3. $$\int N(\xi, {}^1T, \mathfrak{J}^0)\, d\xi = \int d\xi^1 \int N((\xi^1, \xi^2), {}^1T, \mathfrak{J}^0)\, d\xi^2.$$

By the result cited above, it follows that $f(u^1, \gamma^2)$ is of bounded variation on $[\alpha^1, \beta^1]$ and relation 1 holds for almost every choice of $u^2=\gamma^2$ in $[\alpha^2, \beta^2]$. In view of relation 3, it follows that $V_{u1}(f; u^2)$ is a summable function of u^2 on $[\alpha^2, \beta^2]$. A similar reasoning applies to the continuous transformation defined by

$$^2T:\quad \xi^1 = u^1, \qquad \xi^2 = f(u^1, u^2), \qquad (u^1, u^2) \text{ in } \mathfrak{J}.$$

These results are summarized in the

LEMMA. *Let $f(u)=f(u^1, u^2)$ be a real, single-valued function defined and continuous on the interval $\mathfrak{J}$. Consider the two transformations*

$$^1T:\quad \xi^1 = u^2, \qquad \xi^2 = f(u^1, u^2), \qquad (u^1, u^2) \text{ in } \mathfrak{J};$$

$$^2T:\quad \xi^1 = u^1, \qquad \xi^2 = f(u^1, u^2), \qquad (u^1, u^2) \text{ in } \mathfrak{J}.$$

A necessary and sufficient condition that $f(u)$ be BV T *on $\mathfrak{J}$ is that both ^1T and ^2T be* BV $\mathfrak{J}^0$. *If $f(u)$ is* BV T *on $\mathfrak{J}$, then*

$$\int N(\xi, {}^1T, \mathfrak{J}^0)d\xi = V_{u1}(f) \geqq \int_{\mathfrak{J}} |f_{u1}(u)|\, du;$$

$$\int N(\xi, {}^2T, \mathfrak{J}^0)d\xi = V_{u2}(f) \geqq \int_{\mathfrak{J}} |f_{u2}(u)|\, du.$$

A necessary and sufficient condition that the sign of equality hold in both these relations is that $f(u)$ be AC T *on $\mathfrak{J}$.*

8. Again let $f(u)$ be a real, single-valued function defined and continuous on the interval $\mathfrak{J}=[\alpha, \beta]$. Retain the notation of the preceding section. For an interval $I=[\gamma^1, \delta^1; \gamma^2, \delta^2]$ contained in $\mathfrak{J}$ define

$$^1\psi(I) = \int_{\gamma^2}^{\delta^2} |f(\delta^1, u^2) - f(\gamma^1, u^2)|\, du^2.$$

It will now be shown that for any interval system $S(\mathfrak{J})$ (see 4),

1. $$^1\psi(S(\mathfrak{J})) \leqq \int \mathrm{K}(\xi, {}^1T, \mathfrak{J}^0)d\xi$$

provided that $\mathrm{K}(\xi^1, T, \mathfrak{J})$ is summable. Evidently $^1\psi(I)$ is the measure of the set of points $\xi=(\xi^1, \xi^2)$ which satisfy the inequalities

$$(f(\delta^1, u^2) - \xi^2)(f(\gamma^1, u^2) - \xi^2) < 0, \qquad \gamma^2 < \xi^1 = u^2 < \delta^2.$$

Consider any point $\xi_0=(\xi_0^1, \xi_0^2)$ which satisfies these inequalities. Its inverse $^1T^{-1}(\xi_0)$ evidently is contained on the line $u^2=\xi_0^1$. Since $f(u^1, \xi_0^1)$ is continuous on $[\gamma^1, \delta^1]$ and has opposite signs at the points $u^1=\gamma^1$, $u^2=\delta^1$, it follows that ξ_0 has a model in the interior of I, and that $\mu(\xi_0, {}^1T, I)=\pm 1$. From II, 6

it follows that $K(\xi_0, {}^1T, I^0) \geqq 1$. Thus ${}^1\psi(I) \leqq \int K(\xi, T, I^0)d\xi$. From II, **19**, **2** it follows that

$$ {}^1\psi(S(\mathfrak{J})) \leqq \sum_{I \text{ in } S(\mathfrak{J})} \int K(\xi, T, I^0)d\xi \leqq \int K(\xi, T, \mathfrak{J}^0)d\xi. $$

Thus relation 1 is established. Now consider the sequence of subdivisions $S_m(\mathfrak{J})$ of $\mathfrak{J}$ into intervals $[\gamma^1_{j-1}, \gamma^1_j; \alpha^2, \beta^2]$, where $\gamma^1_j = \alpha^1 + (\beta^1 - \alpha^1)j/m$, $j = 0, \cdots, m$. Evidently

$$ {}^1\psi(S_m(\mathfrak{J})) = \int_{\alpha^2}^{\beta^2} \sum_{j=1}^{m} \left| f(\gamma^1_j, u^2) - f(\gamma^1_{j-1}; u^2) \right| du^2 $$

and (see III, **5**)

$$ \lim \sum_{j=1}^{m} \left| f(\gamma^1_j, u^2) - f(\gamma^1_{j-1}, u^2) \right| = V_{u^1}(f; u^2). $$

From the lemma of Fatou it follows that $V_{u^1}(f; u^2)$ is summable on $[\alpha^2, \beta^2]$ whenever $K(\xi, {}^1T, \mathfrak{J}^0)$ is summable, and

$$ V_{u^1}(f) \leqq \int K(\xi, {}^1T, \mathfrak{J}^0)d\xi. $$

Since $K(\xi, {}^1T, \mathfrak{J}^0) \leqq \bar{N}(\xi, {}^1T, \mathfrak{J}^0)$, it is clear from III, **7** that a necessary and sufficient condition that $V_{u^1}(f; u^2)$ be summable on $[\alpha_2, \beta_2]$ is that $K(\xi, {}^1T, \mathfrak{J}^0)$ be summable. If $K(\xi, {}^1T, \mathfrak{J}^0)$ is summable, then

$$ V_{u^1}(f) = \int K(\xi, {}^1T, \mathfrak{J}^0)d\xi \geqq \int_{\mathfrak{J}} \left| f_{u^1}(u) \right| du. $$

Similar statements are valid for the transformation ${}^2T$. Combining these facts with those in III, **3**, **4**, **7**, one obtains the

LEMMA. *Let $f(u) = f(u^1, u^2)$ be a real, single-valued function defined and continuous on the interval $\mathfrak{J}$. Consider the flat surfaces having the representations*

$$ {}^1S: \quad [(u^2, f(u^1, u^2)), \mathfrak{J}]; \qquad {}^2S: \quad [(u^1, f(u^1, u^2)), \mathfrak{J}]. $$

A necessary and sufficient condition that $f(u)$ be BV T *on $\mathfrak{J}$ is that both ${}^1S$ and ${}^2S$ be* BEV. *If $f(u)$ is* BV T *on $\mathfrak{J}$, then*

$$ eV({}^1S) = V_{u^1}(f) \geqq \int_{\mathfrak{J}} \left| f_{u^1}(u) \right| du; \qquad eV({}^2S) = V_{u^2}(f) \geqq \int_{\mathfrak{J}} \left| f_{u^2}(u) \right| du. $$

A necessary and sufficient condition that the sign of equality hold in both these relations is that $f(u)$ be AC T *on $\mathfrak{J}$. A necessary and sufficient condition that $f(u)$ be* AC T *on $\mathfrak{J}$ is that both of the representations for ${}^1S$ and ${}^2S$ be* AC E.

Chapter IV

On the essential area

1. Let S be a continuous surface in x-space (see I, 1), and let $[x(u), \mathfrak{B}]$ be any representation for S. With S there is associated the three projection surfaces ${}^iS$ upon the coordinate planes ${}^ix$ having representations $[{}^ix(u), \mathfrak{B}]$, which determine bounded continuous transformations ${}^iT$ for $i=1, 2, 3$ (see **4**; I, **4**). In this chapter, the theory of the preceding chapters is used when the plane ξ coincides in turn with the coordinate planes ${}^ix$, $i=1, 2, 3$. The following triple notation is useful (see III, **2**)

$$\mathrm{K}(x, S) \equiv (\mathrm{K}({}^1x, {}^1S), \mathrm{K}({}^2x, {}^2S), \mathrm{K}({}^3x, {}^3S));$$

$$\mathcal{J}(u, [x, \mathfrak{B}]) \equiv (\mathcal{J}(u, {}^1T), \mathcal{J}(u, {}^2T), \mathcal{J}(u, {}^3T)) \quad \text{for} \quad u \text{ in } \mathfrak{B}^0;$$

$$J(u, [x, \mathfrak{B}]) \equiv (J(u, {}^1T), J(u, {}^2T), J(u, {}^3T)) \quad \text{for} \quad u \text{ in } \mathfrak{B}^0.$$

The triple $[x(u), \mathfrak{B}]$ is said to be BV E, AC E, and so on, when each of the associated transformations ${}^iT$ is BV $E({}^iT, \mathfrak{B}^0)$, AC $E({}^iT, \mathfrak{B}^0)$, and so on, for $i=1, 2, 3$ (cf. III, **1**).

2. Given a continuous surface S for which $\mathrm{K}(x, S)$ is summable, that is, for which each of the projection surfaces ${}^1S$, ${}^2S$, ${}^3S$ is BEV (see III, **3**). If $[x(u), \mathfrak{B}]$ be any representation for S, define, for any simple Jordan region B in $\mathfrak{B}$

$$ {}^i\phi(B) = \int \mathrm{K}({}^ix, {}^iT, B^0)d^ix, \qquad i = 1, 2, 3;$$

$$\phi(B) = ({}^1\phi(B), {}^2\phi(B), {}^3\phi(B)), \qquad \Phi(B) = \|\phi(B)\| \quad \text{for} \quad B \text{ in } \mathfrak{B}.$$

It follows from II, **19** that, for any finite system $S(\mathfrak{B})$ of nonoverlapping simple Jordan regions in $\mathfrak{B}$ (see **4**), it is true that ${}^i\phi(S(\mathfrak{B})) \leqq {}^i\phi(\mathfrak{B})$ for $i=1, 2, 3$. Thus one has (see III **1–3**)

$$eV({}^iS) = {}^i\phi(\mathfrak{B}) = U(\mathfrak{B}; [{}^i\phi, \mathfrak{B}]) \qquad \text{for } i = 1, 2, 3. \tag{1}$$

From a remark in **5** follows the fact that $U(\mathfrak{B}; [\Phi, \mathfrak{B}])$ is finite; moreover, one has the

Lemma. *The quantity $U(\mathfrak{B}; [\Phi, \mathfrak{B}])$ is independent of the choice of the representation for S.*

Proof. Let $[x_1(u), \mathfrak{B}_1]$, $[x_2(u), \mathfrak{B}_2]$ be any two representations for S. From I, **1** it follows that there exists a sequence of topological maps of $\mathfrak{B}_1$ onto $\mathfrak{B}_2$ given by $[\bar{u}_n(u), \mathfrak{B}_1]$ for $n=3, 4, \cdots$ such that $\|x_1(u)-x_2(\bar{u}_n(u))\| < n^{-1}$ for u in $\mathfrak{B}_1$. Denote by ${}^iT_n$ the transformations given by the triples $[{}^ix_2(\bar{u}_n(u)), \mathfrak{B}_1]$ for $n=3, 4, \cdots$; $i=1, 2, 3$. Put

$$ {}^i\phi_n(B) = \int \mathrm{K}({}^ix, {}^iT_n, B^0)d^ix, \qquad i = 1, 2, 3;\ n = 1, 2, 3, \cdots;$$

$$\phi_n(B) = ({}^1\phi_n(B), {}^2\phi_n(B), {}^3\phi_n(B)), \qquad \Phi_n(B) = \|\phi_n(B)\| \quad \text{for} \quad B \text{ in } \mathfrak{B}.$$

To any finite system $S(\mathfrak{B})$ (see **4**) there corresponds under the map $[\bar{u}_n(u), \mathfrak{B}_1]$ a finite system $S_n(\mathfrak{B}_2)$. Clearly (see III, **1**)

$$U(\mathfrak{B}_2, [\Phi_2, \mathfrak{B}_2]) \geqq \Phi_2(S_n(\mathfrak{B}_2)) = \Phi_n(S(\mathfrak{B}_1)) \qquad \text{for } n = 3, 4, \cdots.$$

Since $\lim \rho({}^iT_n, {}^iT_1; \mathfrak{B}_1) = 0$ for $i = 1, 2, 3$, it follows from II, **2**, 2 and the lemma of Fatou that $\lim\inf \Phi_n(S(\mathfrak{B}_1)) \geqq \Phi_1(S(\mathfrak{B}_1))$. From the preceding relations and the definition of the U-function (see **4**), it follows that $U(\mathfrak{B}_2; [\Phi_2, \mathfrak{B}_2]) \geqq U(\mathfrak{B}_1; [\Phi_1, \mathfrak{B}_1])$. The opposite inequality follows by symmetry, and the lemma is established.

3. Let S be a continuous surface in x-space. In view of the preceding lemma, one may define the *essential area* $eA(S)$ for the surface as follows. If $[x(u), \mathfrak{B}]$ be any representation for S, set

$$eA(S) = \begin{cases} U(\mathfrak{B}; [\Phi, \mathfrak{B}]) & \text{if } \mathrm{K}(x, S) \text{ is summable;} \\ +\infty & \text{otherwise.} \end{cases}$$

Clearly $eA(S)$ is independent of the choice of the representation for S, although it is not clear as to whether it is also independent of the choice of a coordinate system in x-space. If S is a flat surface in a coordinate plane (see I, **3**), then $eA(S) = eV(S)$ (see III, **2**; IV, **2**, 1)[8].

4. From **5**; IV, **2**, 1 follows the

THEOREM. *A necessary and sufficient condition that the essential area $eA(S)$ of a surface S be finite is that each of the projection surfaces ${}^1S, {}^2S, {}^3S$ be of bounded essential variation. Between the essential area and the essential variations of the projection surfaces, the following relation exists*

$$eV({}^iS) \leqq eA(S) \leqq \sum_{i=1}^{3} eV({}^iS).$$

Notice that this theorem is an analogue for continuous surfaces of the result for continuous curves cited in **1**, 1.

5. LEMMA. *The essential area $eA(S)$ is a lower semi-continuous functional of S.*

Proof. Suppose the sequence of continuous surfaces S_n converges to a surface S_0; from I, **1** it follows that there exist representations $[x_n(u), \mathfrak{B}]$ for $n = 0, 1, 2, \cdots$ such that $x_n(u)$ converges on $\mathfrak{B}$ uniformly to $x_0(u)$. Adopt the notation of IV, **2**, using a subscript n to distinguish the functions belonging to

---

(8) Since $eV({}^iS) = eA({}^iS)$ for $i = 1, 2, 3$, the notion of essential variation might very well be discarded. However, it has not been the custom to speak of the length of a one-dimensional curve, but rather to speak of the total variation of a function representing that curve. To preserve this parallel between the theory of curves and the theory of surfaces, the concept of essential variation has been introduced here.

S_n. From II, 2, 2 and the lemma of Fatou, it follows that $\liminf {}^i\phi_n(B) \geqq {}^i\phi_0(B)$ for B in $\mathfrak{B}$, $i=1, 2, 3$. From the lemma in **5** and the definition of the essential area follows $\liminf eA(S_n) \geqq eA(S_0)$, and the lemma is proved.

6. THEOREM. *If the essential area for a surface S is finite, and if $[x(u), \mathfrak{B}]$ be any representation for S, then the triple $\mathcal{J}(u, [x, \mathfrak{B}])$ of generalized Jacobians exists almost everywhere in $\mathfrak{B}^0$, is summable on $\mathfrak{B}^0$, and*

1. $$eA(S) \geqq \int_{\mathfrak{B}^0} \|\mathcal{J}(u, [x, \mathfrak{B}])\| du.$$

A sufficient condition that the sign of equality hold here is that the triple $[x(u), \mathfrak{B}]$ be AC E; *a necessary condition that the sign of equality hold is that the triple be* AC $\mathcal{E}$.

Proof. Since $eA(S)$ is finite, it is true that each ${}^iS$ is BEV (see IV, **4**), hence $K({}^ix, {}^iT, \mathfrak{B}^0)$ is summable for $i=1, 2, 3$ (see III, **2–3**). Thus (see IV, **2**; II, **9**, 1) each $\mathcal{J}(u, {}^iT)$ exists almost everywhere in $\mathfrak{B}^0$, is summable on $\mathfrak{B}^0$, and

2. $${}^i\phi(B) \geqq \int_{B^0} | \mathcal{J}(u, {}^iT) | du \quad \text{for} \quad B \text{ in } \mathfrak{B}, \qquad i = 1, 2, 3.$$

Define

3. $${}^i\psi(B) = \int_{B^0} | \mathcal{J}(u, {}^iT) | du, \qquad i = 1, 2, 3;$$
$$\psi(B) = ({}^1\psi(B), {}^2\psi(B), {}^3\psi(B)), \qquad \Psi(B) = \|\psi(B)\| \quad \text{for} \quad B \text{ in } \mathfrak{B}.$$

If $S(\mathfrak{B})$ be any finite system, it follows by a known inequality (see Hardy, Littlewood, Polya [1, chap. VI]) that

$$\Psi(S(\mathfrak{B})) \leqq \int_{\mathfrak{B}^0} \|\mathcal{J}(u, [x, \mathfrak{B}])\| du.$$

Thus

4. $$U(\mathfrak{B}; [\Psi, \mathfrak{B}]) \leqq \int_{\mathfrak{B}^0} \|\mathcal{J}(u, [x, \mathfrak{B}])\| du.$$

On the other hand, let R be a simple Jordan region in $\mathfrak{B}$ which admits of an interval subdivision. Then it is true (see R[2] [2, chap. II, §10]) that there exists a sequence of interval subdivisions $S_n(R)$ such that

$$\lim \Psi(S_n(R)) = \int_{R^0} \|\mathcal{J}(u, [x, \mathfrak{B}])\| du.$$

Thus

5. $$U(\mathfrak{B}; [\Psi, \mathfrak{B}]) \geqq \int_{R^0} \|\mathcal{J}(u, [x, \mathfrak{B}])\| du.$$

Since the open domain $\mathfrak{B}^0$ may be filled up from the interior (see **4**) by a sequence of simple regions of type R (see Kerékjártó [1]), one concludes from relations 4 and 5 that

6. $$U(\mathfrak{B};\ [\Psi,\mathfrak{B}]) = \int_{\mathfrak{B}^0} \|\mathcal{J}(u,\ [x,\mathfrak{B}])\| du.$$

Relations 2 and 3 imply

7. $$U(\mathfrak{B};\ [\Phi,\mathfrak{B}]) \geqq U(\mathfrak{B};\ [\Psi,\mathfrak{B}]).$$

In view of IV, **3**, the first statement in the theorem is established. Now (see II, **9**, **12**) if $[x(u),\mathfrak{B}]$ is AC E, then the sign of equality holds in 2, hence in 7. Thus a sufficient condition that the sign of equality hold in 1 is that $[x(u),\mathfrak{B}]$ be AC E. Next, suppose the sign of equality holds in relation 1. Let I be any interval in $\mathfrak{B}$; extend the lines forming the boundary of I indefinitely. The simple Jordan region $\mathfrak{B}$ is thus divided into a possibly enumerable number of simple Jordan regions $I \equiv B_0, B_1, B_2, \cdots$. For any simple Jordan region B in $\mathfrak{B}$, the reasoning leading to relation 6 gives

$$U(B;\ [\Psi,\mathfrak{B}]) = \int_{B^0} \|\mathcal{J}(u,\ [x,\mathfrak{B}])\| du.$$

Since a bounded portion of a straight line is rectifiable, it follows that

$$U(\mathfrak{B};\ [\Psi,\mathfrak{B}]) = \sum_n U(B_n;\ [\Psi,\mathfrak{B}]).$$

From this relation and relations 2, 3, one finds

$$U(\mathfrak{B};[\Phi,\mathfrak{B}]) \geqq \sum_n U(B_n;\ [\Phi,\mathfrak{B}]) \geqq \sum_n U(B_n;\ [\Psi,\mathfrak{B}]) = U(\mathfrak{B};\ [\Psi,\mathfrak{B}]).$$

But since the sign of equality holds in 1 by hypothesis, it follows that the sign of equality must hold at every step in the preceding inequalities. In particular, then

$$U(I;\ [\Phi,\mathfrak{B}]) = U(I;\ [\Psi,\mathfrak{B}]) = \int_I \|\mathcal{J}(u,\ [x,\mathfrak{B}])\| du.$$

If s be any closed oriented square in $\mathfrak{B}$, then obviously (see II, **3–8**; IV, **1–3**)

$$\mathcal{G}(s,\ {}^iT,\ \mathcal{E}({}^iT,\mathfrak{B}^0)) \leqq \int \mathrm{K}({}^ix,\ {}^iT,\ s^0)d^ix \leqq U(s;\ [\Phi,\mathfrak{B}])$$

$$= \int_s \|\mathcal{J}(u,\ [x,\mathfrak{B}])\| du.$$

Thus each ${}^iT$ is AC $\mathcal{E}({}^iT,\ \mathfrak{B}^0)$ for $i=1, 2, 3$—that is, $[x(u),\mathfrak{B}]$ is AC $\mathcal{E}$. This completes the proof.

7. Corollary. *If a surface S has a representation $[x(u), \mathfrak{B}]$ which is* BV E, *then the essential area $eA(S)$ is finite, the triple $\mathcal{J}(u, [x, \mathfrak{B}])$ of generalized Jacobians exists almost everywhere in $\mathfrak{B}^0$, is summable on $\mathfrak{B}^0$, and*

$$eA(S) \geqq \int_{\mathfrak{B}^0} \|\mathcal{J}(u, [x, \mathfrak{B}])\| du.$$

A necessary and sufficient condition that the sign of equality hold here is that $[x(u), \mathfrak{B}]$ be AC E.

A proof follows from results in II, 11; III, 3, 1; IV, 4, 6. This is an analogue to the theorems in 1, 2, 3 for continuous curves; all representations for continuous curves automatically satisfy the analogue of BV E in one dimension whenever the length of the curves is finite.

8. Corollary. *Let S_n, $n=0, 1, 2, \cdots$, be a sequence of continuous surfaces satisfying the following conditions: the surfaces S_n converge to S_0 (see* I, 1); *each of the surfaces has a finite essential area $eA(S_n)$ for $n=0, 1, 2, \cdots$ (see* IV, 3, 4); *the surface S_0 has a representation $[x(u), \mathfrak{B}]$ for which the essential areas $eA(S_n)$ converge to $\int_{\mathfrak{B}^0}\|\mathcal{J}(u, [x, \mathfrak{B}])\|du$. Then the representation $[x(u), \mathfrak{B}]$ is* AC $\mathcal{E}$.

Proof. From IV, 5, 6 and the assumptions, one obtains

$$\int_{\mathfrak{B}^0} \|\mathcal{J}(u, [x, \mathfrak{B}])\| du = \lim eA(S_n) \geqq eA(S_0) \geqq \int_{\mathfrak{B}^0} \|\mathcal{J}(u, [x, \mathfrak{B}])\| du.$$

Thus the sign of equality holds throughout, and the conclusion now follows from the last part of the theorem in IV, 6.

9. By the principle stated in I, 8, the theorem in IV, 6 may be given the following variant form.

Theorem. *A necessary condition that a representation $[x(u), \mathfrak{B}]$ for a surface S be absolutely continuous $(eA, \mathcal{J})$, where $eA(S)$ is the essential area of S and $\mathcal{J}(u, [x, \mathfrak{B}])$ is the triple of generalized Jacobians, is that $[x(u), \mathfrak{B}]$ be* AC $\mathcal{E}$. *A sufficient condition that this representation for S be absolutely continuous $(eA, \mathcal{J})$ is that it be* AC E; *then each of the corresponding representations for the projection surfaces ${}^iS$ is also absolutely continuous $(eA, \mathcal{J})$ for $i=1, 2, 3$.*

10. The results in IV, 6–9 are paralleled by similar theorems involving the essential area and the triple of ordinary Jacobians; one need but replace II, 9, 1 by II, 9, 2 in making the proofs.

Theorem. *If the essential area for a surface S is finite, and if $[x(u), \mathfrak{B}]$ be any representation for S for which the triple $J(u, [x, \mathfrak{B}])$ of ordinary Jacobians exists almost everywhere in $\mathfrak{B}^0$, then the triple is summable on $\mathfrak{B}^0$, and*

$$eA(S) \geqq \int_{\mathfrak{B}^0} \|J(u, [x, \mathfrak{B}])\| du.$$

A sufficient condition that the sign of equality hold here is that the triple $[x(u)\ \mathfrak{B}]$ *be* AC E; *a necessary condition that the sign of equality hold is that the triple be* AC $\mathcal{E}$.

11. COROLLARY. *If a surface S has a representation $[x(u),\ \mathfrak{B}]$ which is* BV E *and for which the triple $J(u,\ [x,\ \mathfrak{B}])$ of ordinary Jacobians exists almost everywhere in $\mathfrak{B}^0$, then the essential area is finite, the triple $J(u,\ [x,\ \mathfrak{B}])$ is summable on $\mathfrak{B}^0$, and*

$$eA(S) \geqq \int_{\mathfrak{B}^0} \| J(u,\ [x,\ \mathfrak{B}]) \| du.$$

A necessary and sufficient condition that the sign of equality hold here is that $[x(u),\ \mathfrak{B}]$ be AC E.

12. COROLLARY. *Let S_n, $n=0, 1, 2, \cdots$, be a sequence of continuous surfaces satisfying the following conditions: the surfaces S_n converge to S_0; each of the surfaces has a finite essential area $eA(S_n)$ for $n=0, 1, 2, \cdots$; the surface S. has a representation $[x(u),\ \mathfrak{B}]$ for which the triple $J(u,\ [x,\ \mathfrak{B}])$ of ordinary Jacobians exists almost everywhere in $\mathfrak{B}^0$ and the essential areas $eA(S)$ converge to $\int_{\mathfrak{B}^0} \| J(u,\ [x,\ \mathfrak{B}]) \| du$. Then the representation $[x(u),\ \mathfrak{B}]$ is* AC $\mathcal{E}$.

13. Here is a variant form for the theorem in IV, **10**.

THEOREM. *A necessary condition that a representation $[x(u),\ \mathfrak{B}]$ for a surface S, for which the triple $J(u,\ [x,\ \mathfrak{B}])$ of ordinary Jacobians exists almost everywhere in $\mathfrak{B}^0$, be absolutely continuous $(eA,\ J)$ is that $[x(u),\ \mathfrak{B}]$ be* AC $\mathcal{E}$. *A sufficient condition that this representation for S be absolutely continuous $(eA,\ J)$ is that it be* AC E; *then each of the corresponding representations for the projection surfaces ${}^iS$ is also absolutely continuous $(eA,\ J)$ for $i=1, 2, 3$.*

14. Rado (see Rado [2]) has shown that a representation $[x(u),\ \mathfrak{B}]$ for a surface S which satisfies a Lipschitz condition of the form $\| x(u_1) - x(u_2) \| \leqq L \| u_1 - u_2 \|$ where L is a constant, is absolutely continuous $(A,\ J)$, where $A(S)$ is the Lebesgue area of S (see I, **6**) and $J(u,\ [x,\ \mathfrak{B}])$ is the triple of ordinary Jacobians. Now the representation $[x(u),\ \mathfrak{B}]$ is also AC E (see II, **8**). From the theorem in IV, **13** one concludes that the essential area and the Lebesgue area of S are equal. In particular, $eA(P) = A(P)$ for every polyhedron P (see I, **5**).

Given any continuous surface S, there exists a sequence of polyhedra P_n such that P_n converge to S and the Lebesgue areas $A(P_n)$ converge to $A(S)$ (see I, **6**, 2). Since $eA(P) = A(P_n)$, one finds by using the lemma in IV, **5** the

THEOREM. *The essential area of a surface S does not exceed the Lebesgue area —that is, $eA(S) \leqq A(S)$. A necessary and sufficient condition that the essential area and the Lebesgue area of a surface S be equal is that there exists a sequence of polyhedra P_n such that P_n converges to S and $eA(P_n)$ converges to $eA(S)$.*

This theorem and the theorem in IV, **10** have the

COROLLARY. *A sufficient condition that the Lebesgue area and the essential area of a surface S be equal is that S possess a representation* $[x(u), \mathfrak{B}]$ *which is absolutely continuous* (A, J), *where* $A(S)$ *is the Lebesgue area of S and* $J(u, [x, \mathfrak{B}])$ *is the triple of ordinary Jacobians.*

15. In order to compare these results with those of Geöcze and Tonelli (cf. **2**), and to give a proof for the statement made in I, **7** the following result is needed.

LEMMA. *If a surface S has a representation* $[x(u), \mathfrak{J}]$ *of non-parametric origin*

$$S:\ x(u) = (u^1, u^2; x^3(u^1, u^2)), \qquad u = (u^1, u^2) \text{ in } \mathfrak{J},$$

then $eA(S)=A_*(S)=A(S)$.

Proof. Define, for any interval $I=[\gamma^1, \delta^1; \gamma^2, \delta^2]$ in $\mathfrak{J}$(⁹).

$$\psi(I) = \left(\int_{\gamma^2}^{\delta^2} | x^3(\gamma^1, u^2) - x^3(\delta^1, u^2) | \, du^2, \right.$$
$$\left. \int_{\gamma^1}^{\delta^1} | x^3(u^1, \gamma^2) - x^3(u^2, \delta^2) | \, du^1, \ | I | \right), \qquad \Psi(I) = \|\psi(I)\|.$$

If ${}^iT$, $i=1, 2, 3$, be the associated continuous transformations (see IV, **1**), then the reader will observe that the transformations ${}^1T$ and ${}^2T$ bear the same relation to the function $x^3(u)$ as those in III, **7**, **8** bear to the function $f(u)$ therein considered; the transformation ${}^3T$ is simply the identity mapping of $\mathfrak{J}$ in the u-plane onto a congruent interval in the ${}^3x$-plane. Define a triple of b-functions $[\phi, \mathfrak{J}]$ as in IV, **2**. The reasoning in III, **8** shows that $\psi(I) \leqq \phi(I)$ for I in $\mathfrak{J}$. Now Rado has shown that (see Rado [1])

$$A_*(S) = \text{l.u.b. } \Psi(S(\mathfrak{J})) \text{ for interval subdivisions } S(\mathfrak{J}).$$

From these relations it is clear that

$$A_*(S) \leqq U(\mathfrak{J}; [\Phi, \mathfrak{J}]) = eA(S).$$

In view of the relations established in I, **7** and IV, **14**, the lemma is now proved.

The reader will notice that attention is restricted here to surfaces having a representation of non-parametric origin where the parameter range is an interval. Further considerations would establish this result more generally, but these are not necessary since it has been customary in the literature on such surfaces to so restrict the range of the independent variables.

16. The lemmas in III, **7**, **8**; IV, **4**, **10**, **11**, **15** imply all the results of

(⁹) These are the expressions of Geöcze (see Rado [1]).

Geöcze and Tonelli for the non-parametric case which are analogous to those for curves stated in **1**. For if S be a surface having a representation $[x(u), \mathfrak{J}]$ of non-parametric origin

$$\text{1.} \qquad x(u) = (u^1, u^2, x^3(u^1, u^2)), \qquad u = (u^1, u^2) \text{ in the interval } \mathfrak{J},$$

then $eA(S) = A(S)$ by the lemma in IV, **15**. Thus from IV, **4** it follows that a necessary and sufficient condition that the Lebesgue area of S be finite is that each of the projection surfaces ${}^1S$, ${}^2S$, ${}^3S$ be BEV. The projection surface ${}^3S$ is obviously BEV under any conditions, while it follows from the lemma in III, **8** that a necessary and sufficient condition that ${}^1S$ and ${}^2S$ be BEV is that $x^3(u)$ be BV T on $\mathfrak{J}$. Summarizing these facts, one obtains a known theorem in the non-parametric case (see Tonelli [1]): a necessary and sufficient condition that the Lebesgue area $A(S)$ of a surface S having a representation **1** of non-parametric origin be finite is that $x^3(u)$ be BV T on $\mathfrak{J}$.

If $x^3(u)$ is BV T on $\mathfrak{J}$, it follows that the triple $J(u, [x, \mathfrak{J}])$ of ordinary Jacobians exists almost everywhere on $\mathfrak{J}$—in fact,

$$J(u, [x, \mathfrak{J}]) = (-\overset{3}{x}_{u^1}(u), \overset{3}{x}_{u^2}(u), 1) \text{ almost everywhere on } \mathfrak{J}.$$

Hence from IV, **10** one concludes that this triple is summable on $\mathfrak{J}$ and

$$A(S) \geqq \int_{\mathfrak{J}^0} \| J(u, [x, \mathfrak{J}]) \| du = \int_{\mathfrak{J}} [\overset{3}{x}_{u^1}(u)^2 + \overset{3}{x}_{u^2}(u)^2 + 1]^{1/2} du.$$

In view of the lemma in III, **7**, it is clear that the representation $[x(u), \mathfrak{J}]$ is BV $\mathfrak{J}^0$, hence BV E, when $x^3(u)$ is BV T on $\mathfrak{J}$; from IV, **11** it follows that a necessary and sufficient condition that the sign of equality hold above is that $[x(u), \mathfrak{J}]$ be AC E. The representation $[{}^3x(u), \mathfrak{J}]$ for ${}^3S$ is clearly AC E under any conditions, while it follows from the lemma in III, **8** that a necessary and sufficient condition that the representations $[{}^1x(u), \mathfrak{J}]$, $[{}^2x(u), \mathfrak{J}]$ be AC E is that $x^3(u)$ be AC T on $\mathfrak{J}$. Thus other known theorems for the non-parametric case are obtained (see Tonelli [1]): if the Lebesgue area $A(S)$ is finite, then each of the partial derivatives

$$\overset{3}{x}_{u^1}(u), \quad \overset{3}{x}_{u^2}(u)$$

exists almost everywhere in $\mathfrak{J}$, is summable on $\mathfrak{J}$ and

$$A(S) \geqq \int_{\mathfrak{J}} [\overset{3}{x}_{u^1}(u)^2 + \overset{3}{x}_{u^2}(u)^2 + 1]^{1/2} du;$$

a necessary and sufficient condition that the sign of equality hold here is that $x^3(u)$ be AC T on $\mathfrak{J}$.

17. The next sections are devoted to a comparison of the essential area defined in IV, **3** with the Geöcze area as defined by Rado (see Rado [2, part II, §1]). Given a continuous surface S, let $[x(u), \mathfrak{B}]$ be any representation for it. Define (see II, **2**, **22**), for any simple Jordan region B in $\mathfrak{B}$,

$$^i\psi(B) = |(\bar{\mathfrak{R}}1, {}^iT, B)| = |{}^iT(E({}^iT, B^0))| = \int c({}^ix, {}^iT, B^0)d\,{}^ix, \qquad i = 1, 2, 3;$$

$$\psi(B) = ({}^1\psi(B), {}^2\psi(B), {}^3\psi(B)), \qquad \Psi(B) = \|\psi(B)\|.$$

The quantity $U(\mathfrak{B}; [\Psi, \mathfrak{B}])$ is independent of the choice of a representation for S (see Rado [2]). It is this quantity which Rado terms the Geöcze area of S—denote it by $\mathcal{G}(S)$. Since $\mathfrak{R}(1, {}^iT, B)$ is the set of points where $K({}^ix, {}^iT, B^0) > 0$, it follows at once that (see IV, **2**, **3**) $\psi(B) \leqq \phi(B)$, hence $\Psi(B) \leqq \Phi(B)$ for B in $\mathfrak{B}$, and $U(\mathfrak{B}, [\Psi, \mathfrak{B}]) \leqq U(\mathfrak{B}, [\Phi, \mathfrak{B}])$. Therefore one has the

LEMMA. *The Geöcze area does not exceed the essential area—that is,* $\mathcal{G}(S) \leqq eA(S)$.

18. Retain the notation of the preceding section. Notice that (see II, **3**, **6**, **7**) for any closed oriented square s in $\mathfrak{B}$, it is so that

$$\Psi(s) \geqq {}^i\psi(s) \geqq |{}^iT(s^0 \cdot \mathcal{E}({}^iT, \mathfrak{B}^0))| = \mathcal{G}(s, {}^iT, \mathcal{E}({}^iT, \mathfrak{B}^0)), \qquad i = 1, 2, 3.$$

Thus a necessary condition that the Geöcze area $\mathcal{G}(S)$ be finite is that each of the representations $[{}^ix(u), \mathfrak{B}]$ for the projection surfaces ${}^iS$ be BV $\mathcal{E}$ for $i = 1, 2, 3$. This result implies the

LEMMA. *If a surface S possesses a representation $[x(u), \mathfrak{B}]$ which is not BV $\mathcal{E}$—that is, for which at least one of the representations $[{}^ix(u), \mathfrak{B}]$ is not BV $\mathcal{E}$—then $\mathcal{G}(S) = eA(S) = +\infty$.*

19. LEMMA. *If a surface S possesses a representation $[x(u), \mathfrak{B}]$ which is BV E, then $\mathcal{G}(S) = eA(S) < +\infty$.*

Proof. From III, **3**; IV, **4** it follows that $eA(S)$ is finite. In view of the lemma in IV, **17** it is sufficient to show that $\mathcal{G}(S) \geqq eA(S)$. Given a positive number ϵ, there exists a finite system $S(\mathfrak{B})$ such that $\Phi(S(\mathfrak{B})) > eA(S) - \epsilon$. Since $[x(u), \mathfrak{B}]$ is BV E, one concludes from the lemma in II, **20** (see IV, **1**) that there exists a sequence of subdivisions $S_n(B)$ for each B in $\mathfrak{B}$ such that $\|S_n(B)\|$ is less than n^{-1} and each $S_n(B)$ is maximal for each of the transformations $[{}^ix_n(u), B]$ for $i = 1, 2, 3$. From the lemma in II, **22** (see IV, **2**, **17**), it follows that ${}^i\phi(B) = \lim {}^i\psi(S_n(B))$ for $i = 1, 2, 3$; B in $S(\mathfrak{B})$. Denote by $S_n(\mathfrak{B})$ the finite system consisting of all simple Jordan regions belonging to an $S_n(B)$ for some B in $S(\mathfrak{B})$. From the triangle inequality one finds that

$$\mathcal{G}(S) \geqq \Psi(S_n(\mathfrak{B})) \geqq \sum_{B \text{ in } S(\mathfrak{B})} \left[\sum_{i=1}^{3} {}^i\psi(S_n(B))^2\right]^{1/2} \to_{n\to\infty} \Phi(S(\mathfrak{B})) > eA(S) - \epsilon.$$

Thus $\mathcal{G}(S) > eA(S) - \epsilon$, and since ϵ is arbitrary, the lemma follows.

20. Summarizing the lemmas in IV, **18**, **19**, one concludes that the essential area and the Geöcze area of a surface are equal if the surface either has a representation which is not BV $\mathcal{E}$, or has a representation which is BV E.

This leaves open the question of whether these areas are equal if all the representations of a surface are BV $\mathcal{E}$, but none is BV E. Indeed, a first question might be whether such surfaces exist. A negative answer to this question would enable one to close the gap between necessary conditions and sufficient conditions in the results in IV, **6**, **10**.

21. For applications in the next chapter, the following result will be useful. Assume that a surface S has a representation which is BV E; in view of the lemma in II, **18**, one may assume that this representation has the form $[x(u), \mathfrak{J}]$ where $\mathfrak{J}$ is an interval. Now the essential area $eA(S)$ is finite (see IV, **4**). It follows from the lemma in II, **20** that there exist interval subdivisions $S_n(\mathfrak{J})$ such that $\|S_n(\mathfrak{J})\|$ converges to zero and each $S_n(\mathfrak{J})$ is maximal for each of the transformations $[{}^i x(u), \mathfrak{J}]$ for $i=1, 2, 3$; for brevity, $S_n(\mathfrak{J})$ is said to be maximal for $[x(u), \mathfrak{J}]$. Given a positive number ϵ, let $S(\mathfrak{J})$ be a finite system of nonoverlapping Jordan regions such that $\Phi(S(\mathfrak{J})) > eA(S) - \epsilon$. For each B in $S(\mathfrak{J})$, denote by $u(B)$ an arbitrary point in the interior of B. For each integer n and each B in $S(\mathfrak{J})$, denote by $S_n(B)$ the maximal collection of those intervals in $S_n(\mathfrak{J})$ whose point set sum β_n is a simple Jordan region containing $u(B)$ and lying in the interior of B. It follows (see Kerékjártó [1]) that the β_n fill up B from the interior (see **4**). Consequently one obtains (see II, **2**, 2; IV, **2**)

$$
\begin{aligned}
{}^i\phi(S_n(B)) &= {}^i\phi(\beta_n) \to_{n\to\infty} {}^i\phi(B) \qquad \text{for } i = 1, 2, 3;\ B \text{ in } S(\mathfrak{J}),\\
eA(S) \geqq \Phi(S_n(\mathfrak{J})) &\geqq \sum_{B \text{ in } S(\mathfrak{J})} \Phi(S_n(B)) \geqq \sum_{B \text{ in } S(\mathfrak{J})} \Phi(\beta_n) \to_{n\to\infty} \Phi(S(\mathfrak{J}))\\
&> eA(S) - \epsilon.
\end{aligned}
$$

In view of IV, **3**, this implies the

LEMMA. *Let S be a surface having a representation which is* BV E. *Then there are representations $[x(u), \mathfrak{J}]$ for S which are* BV E, *and for every sequence of interval subdivisions $S_n(\mathfrak{J})$ such that $\|S_n(\mathfrak{J})\|$ converges to zero and each $S_n(\mathfrak{J})$ is maximal with respect to $[x(u), \mathfrak{B}]$, it is true that* $\lim \Phi(S_n(\mathfrak{J})) = eA(S)$.

CHAPTER V

APPLICATIONS

1. In a study of convergence in area, Rado and Reichelderfer (see R² [2]) obtained the following results. Let S_n, $n=0, 1, 2, \cdots$, be a sequence of continuous surfaces having representations $[x_n(u), \mathfrak{J}]$ of non-parametric origin (see I, **2**), where

$$
x_n(u) = (\overset{1}{u}, \overset{2}{u}, \overset{3}{x}_n(\overset{1}{u}, \overset{2}{u})) \quad \text{for} \quad u = (\overset{1}{u}, \overset{2}{u}) \text{ in } \mathfrak{J}, \qquad n = 0, 1, 2, \cdots.
$$

Make the following assumptions: the functions $\overset{3}{x}_n(u)$ converge on $\mathfrak{J}$ uniformly to $\overset{3}{x}_0(u)$; each of the functions $\overset{3}{x}_n(u)$ for $n=0, 1, 2, \cdots$ is BV T on $\mathfrak{J}$

(see III, **5**); and the Lebesgue areas $A(S_n)$ converge to $A(S_0)$ (see I, **6**). Then the total variations $V_{u^1}(x_n^3)$, $V_{u^2}(x_n^3)$ converge to $V_{u^1}(x_0^3)$, $V_{u^2}(x_0^3)$, respectively (see R[2] [2, chap. I, §19]). Observe that the first assumption implies that S_n converges to S_0 (see I, **1**).

2. Suppose that S is a continuous surface having a representation $[x(u), \mathfrak{J}]$ for which the triple of ordinary Jacobians $J(u; [x, \mathfrak{J}])$ exists almost everywhere in $\mathfrak{J}$ (see IV, **1**) and the Lebesgue area $A(S)$ is finite; then the triple $J(u, [x, \mathfrak{J}])$ is summable on $\mathfrak{J}$ (see IV, **10**, **14**). Assume that S_n, $n=1, 2, \cdots$, is a sequence of continuous surfaces having representations $[x_n(u), \mathfrak{J}]$ satisfying the following conditions: the functions $x_n^i(u)$ converge on $\mathfrak{J}$ uniformly to $x^i(u)$ for $i=1, 2, 3$; each of the representations $[x_n(u), \mathfrak{J}]$ for $n=1, 2, \cdots$, is absolutely continuous (A, J) (see I, **8**); and the Lebesgue areas $A(S_n)$ converge to $\int_{\mathfrak{J}} \| J(u, [x, \mathfrak{J}]) \| du$. Then Rado and Reichelderfer (see R[2] [2, chap. I, §§25–27]) show that

$$\lim \int_{\mathfrak{J}} | J(u, [{}^ix_n, \mathfrak{J}]) | \, du = \int_{\mathfrak{J}} | J(u, [{}^ix, \mathfrak{J}]) | \, du \qquad \text{for } i = 1, 2, 3.$$

Since sequences of surfaces S_n having the properties just described exist if and only if the representation $[x(u), \mathfrak{J}]$ for S is absolutely continuous (A, J), they show that if $[x(u), \mathfrak{J}]$ is absolutely continuous (A, J), then each of the representations $[{}^ix(u), \mathfrak{J}]$ for the projection surfaces ${}^iS$ for $i=1, 2, 3$ is also absolutely continuous (A, J), and each of the transformations $[{}^ix(u), \mathfrak{J}^0]$ for $i=1, 2, 3$ belongs to the class K_3 described in II, **8**. Observe that the first condition on the $[x_n(u), \mathfrak{J}]$ implies that S_n converges to S_0. These results of Rado and Reichelderfer will appear as corollaries to more general theorems whch are presented in the wake of certain preliminary notions (see V, **8**, **15**).

3. Let $\mathfrak{J}$ be any interval in the u-plane. A class $\mathfrak{F}$ of intervals I in $\mathfrak{J}$ is termed closed if it possesses the following properties.

1. The interval $\mathfrak{J}$ is in $\mathfrak{F}$.
2 If I_1 and I_2 are in $\mathfrak{F}$, and if $I_1 \cdot I_2$ is an interval, then $I_1 \cdot I_2$ is in $\mathfrak{F}$.
3. If I is any interval in $\mathfrak{F}$, then there exists an interval subdivision $S(\mathfrak{J})$ in $\mathfrak{F}$ which contains I as an element.

A particular type of closed class is important for the sequel. Let $\mathcal{L}$ be a class containing an at most enumerable number of lines(10), each of which is parallel to one of the coordinate axes in the u-plane, and none of which forms a side of the interval $\mathfrak{J}$. Denote by $\mathfrak{F}(\mathcal{L})$ the class of all intervals I in $\mathfrak{J}$, each of whose four sides is formed by a segment of a line not in $\mathcal{L}$. It is readily verified that $\mathfrak{F}(\mathcal{L})$ is a closed class. If $\mathcal{L}_1, \mathcal{L}_2, \cdots$ be a finite or enumerably infinite set of classes of lines each of type $\mathcal{L}$, then the class $\mathfrak{F}$ of all the intervals found in every one of the classes $\mathfrak{F}(\mathcal{L}_1)$, $\mathfrak{F}(\mathcal{L}_2)$, $\cdots$ is again a closed class having the same structure as $\mathfrak{F}(\mathcal{L})$. A closed class of type $\mathfrak{F}(\mathcal{L})$ is termed c-closed.

(10) The class $\mathcal{L}$ may be empty.

Let $\mathfrak{F}$ be any closed class. Assume that to every interval I in $\mathfrak{F}$ there is associated a finite real number $\phi(I)$; this function is termed an interval function of $\mathfrak{F}$, and denoted by $[\phi, \mathfrak{F}]$. For every interval I in $\mathfrak{F}$, define (cf. **4**)

$$u(I; [\phi, \mathfrak{F}]) = \text{l.u.b. } \phi(S(I)) \text{ for finite interval systems } S(I) \text{ in } \mathfrak{F}.$$

Evidently $\phi(I) \leqq u(I; [\phi, \mathfrak{F}]) \leqq u(\mathfrak{J}; [\phi, \mathfrak{F}])$ for I in $\mathfrak{F}$. A necessary and sufficient condition that $[u, \mathfrak{F}]$ be an interval function on $\mathfrak{F}$ is that $u(\mathfrak{J}; [\phi, \mathfrak{F}])$ be finite; if $[u, \mathfrak{F}]$ is an interval function, then $[\phi, \mathfrak{F}]$ is said to possess a u-function.

4. Let S be a continuous surface having a representation which is BV E; as noted in IV, **21**, one may without loss of generality assume this representation to be of the form $[x(u), \mathfrak{J}]$, where $\mathfrak{J}$ is an interval. Let $\mathfrak{F}([x, \mathfrak{J}])$ denote the class of all intervals I in $\mathfrak{J}$ having the following property: for each of the four lines l, a segment of which forms the boundary of I, it is true that (cf. IV, 1) $|{}^iT(l \cdot E({}^iT, \mathfrak{J}^0))| = 0$ for $i = 1, 2, 3$ where ${}^iT$ is the transformation $[{}^ix(u), \mathfrak{J}^0]$. From the corollary in II, **17**, it is clear that $\mathfrak{F}([x, \mathfrak{J}])$ is a c-closed class. Let I be any interval in $\mathfrak{F}([x, \mathfrak{J}])$ and let $S(I)$ be any interval subdivision in $\mathfrak{F}([x, \mathfrak{J}])$. Then (see II, **19**; IV, **2**, **21**) $S(I)$ is a maximal system for $[x(u), \mathfrak{J}]$, so that

1. $$ {}^i\phi(S(I)) = {}^i\phi(I) \quad \text{for} \quad i = 1, 2, 3; \qquad \Phi(S(I)) \geqq \Phi(I).$$

Let $\mathfrak{F}$ be any c-closed subset of $\mathfrak{F}([x, \mathfrak{J}])$. Then for every interval I in $\mathfrak{F}$, there necessarily exists a sequence of interval subdivisions $S_n(I)$ in $\mathfrak{F}$ such that $\|S_n(I)\|$ converges to zero. From IV, **2**, **21**, one finds (see **4**; V, **3**), for I in $\mathfrak{F}$,

2. $$\begin{aligned} &U(I; [{}^i\phi, \mathfrak{J}]) = u(I; [{}^i\phi, \mathfrak{F}]) = {}^i\phi(I) \qquad \text{for } i = 1, 2, 3; \\ &U(I; [\Phi, \mathfrak{J}]) = u(I; [\Phi, \mathfrak{F}]). \end{aligned}$$

5. General lemma. *Let* $[\phi_n, \mathfrak{F}]$ *where* $\phi_n(I) = ({}^1\phi_n(I), {}^2\phi_n(I), {}^3\phi_n(I))$ *for* I *in* $\mathfrak{F}$, $n = 0, 1, 2, \cdots$, *be a sequence of triples of interval functions defined on a closed class* $\mathfrak{F}$. *Set* $\Phi_n(I) = \|\phi_n(I)\|$ *for* I *in* $\mathfrak{F}$. *Make the following assumptions.*

1. *Each* $[{}^i\phi_n, \mathfrak{F}]$ *is non-negative for* $i = 1, 2, 3$; $n = 0, 1, 2, \cdots$.
2. *Each* $[{}^i\phi_n, \mathfrak{F}]$ *has a u-function for* $i = 1, 2, 3$; $n = 0, 1, 2, \cdots$.
3. *If* I *be any interval in* $\mathfrak{F}$, *if* $S(I)$ *be any interval subdivision in* $\mathfrak{F}$, *then* ${}^i\phi_n(S(I)) \geqq {}^i\phi_n(I)$ *for* $i = 1, 2, 3$; $n = 0, 1, 2, \cdots$.
4. $\liminf {}^i\phi_n(I) \geqq {}^i\phi_0(I)$ *for* I *in* $\mathfrak{F}$; $i = 1, 2, 3$.
5. $\lim u(\mathfrak{J}; [\Phi_n, \mathfrak{F}]) = u(\mathfrak{J}; [\Phi_0, \mathfrak{F}])$.

Then $\lim u(I, [{}^i\phi_n, \mathfrak{F}]) = u(I; [{}^i\phi_0, \mathfrak{F}])$ *for* I *in* $\mathfrak{F}$, $i = 1, 2, 3$.

For the special case when $\mathfrak{F}$ consists of all the intervals in $\mathfrak{J}$, this lemma is stated and proved by Rado and Reichelderfer (see R[2] [2, chap. II, §§1–8]). A proof for this slightly more general lemma may be made by using the properties of a closed class (see V, **3**), and following step by step their proof.

6. A central result in this chapter may now be stated and proved.

THEOREM. *Let S_n, $n=0, 1, 2, \cdots$, be a sequence of continuous surfaces satisfying the following conditions.*

1. *The surfaces S_n converge to S_0 (see* I, 1).
2. *Each of the surfaces S_n for $n=0, 1, 2, \cdots$ has a representation which is* BV E (*see* IV, 1).
3. *The essential areas $eA(S_n)$ converge to $eA(S_0)$.*

Then the essential variations $eV({}^iS_n)$ for the projection surfaces ${}^iS_n$ converge to $eV({}^iS_0)$ for $i=1, 2, 3$.

Proof. From conditions 1, 2; IV, 21; I, 1, it follows that there exist representations $[x_n(u), \mathfrak{J}]$ for S_n each having the same interval of definition $\mathfrak{J}$, and each BV E, such that $x_n(u)$ converges uniformly on $\mathfrak{J}$ to $x_0(u)$. Let $\mathfrak{F}$ denote the class of all intervals belonging to every one of the c-closed classes, $\mathfrak{F}([x_n, \mathfrak{J}])$ for $n=0, 1, 2, \cdots$ (see V, 4); then $\mathfrak{F}$ is a c-closed class. Define triples of b-functions $[\phi_n, \mathfrak{J}]$ as in IV, 2 for $n=0, 1, 2, \cdots$. From V, 4, 2, one obtains, for I in $\mathfrak{F}$ (see III, 3; IV, 3), $n=0, 1, 2, \cdots$,

4.
$$
\begin{aligned}
U(I; [{}^i\phi_n, \mathfrak{J}]) &= u(I; [{}^i\phi_n, \mathfrak{F}]) = {}^i\phi_n(I) \qquad \text{for } i = 1, 2, 3;\\
U(I; [\Phi_n, \mathfrak{J}]) &= u(I; [\Phi_n, \mathfrak{F}]);\\
eV({}^iS_n) &= u(\mathfrak{J}; [{}^i\phi_n, \mathfrak{F}]) = {}^i\phi_n(\mathfrak{J}) \qquad \text{for } i = 1, 2, 3;\\
eA(S_n) &= u(\mathfrak{J}, [\Phi_n, \mathfrak{F}]).
\end{aligned}
$$

Thus conditions V, 5, 1, 2 are satisfied. From V, 4, 1, it is clear that V, 5, 3 is fulfilled. Condition V, 5, 4 follows at once from II, 2, 2 and the lemma of Fatou, since $x_n(u)$ converges on $\mathfrak{J}$ uniformly to $x_0(u)$. Finally, from relation 4 and condition 3 follows condition V, 5, 5. The conclusion of this theorem thus follows at once from the lemma in V, 5 and relation 4.

7. COROLLARY. *Let S be a continuous surface having a representation which is* BV E. *If the Lebesgue area $A(S)$ equals the essential area $eA(S)$, then the Lebesgue areas $A({}^iS)$ of the projection surfaces ${}^iS$ equal the essential areas $eA({}^iS)$ for $i=1, 2, 3$.*

Proof. Let P_n, $n=1, 2, \cdots$, be a sequence of polyhedra such that P_n converges to S and $A(P_n)$ converges to $A(S)$ (see I, 6, 2). Since $A(P_n)=eA(P_n)$ and each P_n has a representation which is AC E (see IV, 14) for $n=1, 2 \cdots$, it follows that the hypotheses of the theorem in V, 6 are fulfilled by P_n and S. Thus $eV({}^iP_n)$ converges to $eV({}^iS)$ for $i=1, 2, 3$. But since ${}^iP_n$ is a flat polyhedron, it is true that $eV({}^iP_n)=eA({}^iP_n)=A({}^iP_n)$ for $i=1, 2, 3$; $n=1, 2, \cdots$ (see IV, 3, 14); also $eV({}^iS)=eA({}^iS)$ for $i=1, 2, 3$. Since ${}^iP_n$ converges to ${}^iS$, it follows that (see I, 6, 3)

$$A({}^iS) \leqq \liminf A({}^iP_n) = \liminf eV({}^iP_n) = eA({}^iS) \qquad \text{for } i = 1, 2, 3.$$

The conclusion of the corollary follows from this inequality and IV, 14.

8. Let S_n, $n=0, 1, 2, \cdots$, be a sequence of continuous surfaces each possessing a representation $[x_n(u), \mathfrak{J}_n]$ of non-parametric origin (see I, **2**), where

$$1. \quad x_n(u) = (\overset{1}{u}, \overset{2}{u}, \overset{3}{x}_n(\overset{1}{u}, \overset{2}{u})) \quad \text{for} \quad u = (\overset{1}{u}, \overset{2}{u}) \text{ in } \mathfrak{J}_n, \quad n = 0, 1, 2, \cdots.$$

Then (see IV, **15**) the essential area $eA(S_n)$ equals the Lebesgue area $A(S_n)$ for $n=0, 1, 2, \cdots$. Suppose that S_n converges to S_0; it follows then that $\mathfrak{J}_n$ converges to $\mathfrak{J}_0$, and $x_n^3(u)$ converges uniformly on every closed set in the interior of $\mathfrak{J}_0$ to $x_0^3(u)$. Now assume that each of the Lebesgue areas $A(S_n)$ is finite for $n=0, 1, 2, \cdots$; this implies (see II, **7**; IV, **16**) that each of the representations $[x_n(u), \mathfrak{J}_n]$ is BV E. Thus to the theorem in V, **6** there follows the

COROLLARY. *Let S_n, $n=0, 1, 2, \cdots$, be a sequence of continuous surfaces, each possessing a representation* 1 *of non-parametric origin and a finite Lebesgue area $A(S_n)$. If the surfaces S_n converge to S_0, if the areas $A(S_n)$ converge to $A(S_0)$, then the variations $eV({}^iS_n)$ for the projection surfaces ${}^iS_n$ converge to $eV({}^iS_0)$ for $i=1, 2, 3$.*

In view of the lemma in III, **8** (see IV, **16**), this result is clearly a generalization of that of Rado and Reichelderfer cited in V, **1**.

9. A second important result in this chapter is contained in the

THEOREM. *Let S_n, $n=0, 1, 2, \cdots$, be a sequence of continuous surfaces satisfying the following conditions:*

1. *the surfaces S_n converge to S_0 (see* I, **1**);
2. *the surface S_0 has a representation $[x_0(u), \mathfrak{B}_0]$ for which the triple $\mathcal{J}(u, [x_0, \mathfrak{B}_0])$ of generalized Jacobians exists almost everywhere in $\mathfrak{B}_0^0$, is summable on $\mathfrak{B}_0^0$ (see* IV, **1**);
3. *the surfaces S_n for $n=1, 2, \cdots$ have representations which are* AC E;
4. *the essential areas $eA(S_n)$ converge to $\int_{\mathfrak{B}_0^0}\|\mathcal{J}(u, [x_0, \mathfrak{B}_0])\|du$.*

Then

5. *the representation $[x_0(u), \mathfrak{B}_0]$ is* AC E;
6. *the representation $[x_0(u), \mathfrak{B}_0]$ is absolutely continuous $(eA, \mathcal{J})$ (see* I, **8**);
7. *the essential variations $eV({}^iS_n)$ for the projection surfaces ${}^iS_n$ converge to $eV({}^iS_0)$ for $i=1, 2, 3$.*

In view of the theorem in IV, **9**, it is clear that conclusion 5 implies conclusion 6; if conclusions 5, 6 are true, then the hypotheses of the theorem in V, **6** are fulfilled, so conclusion 7 follows. It suffices therefore to prove 5. This proof is divided into two parts: an "assume without loss of generality" section (V, **10**), and the proof itself (V, **11**).

10. No loss of generality is imposed in the preceding theorem if the following additional assumptions are made:

1. the simple Jordan regions $\mathfrak{B}_n$ fill up $\mathfrak{B}_0$ from the interior (see **4**);

2. on every closed set in the interior of $\mathfrak{B}_0$, $x_n(u)$ converges uniformly to $x_0(u)$.

In proving this, a theorem of Franklin and Wiener is useful (see Franklin and Wiener [1]). Given a topological map $[\bar{u}(u), \mathfrak{B}]$ of a Jordan region $\mathfrak{B}$ in the u-plane onto a Jordan region $\overline{\mathfrak{B}}$ and a positive constant ϵ, there exists a pair of analytic functions $[\bar{u}_\epsilon(u), \mathfrak{R}]$ defining a topological map of some Jordan region $\mathfrak{R}$ in the u-plane containing $\mathfrak{B}$ in its interior onto a Jordan region $\overline{\mathfrak{R}}$ containing $\overline{\mathfrak{B}}$ in its interior, and such that $\|\bar{u}_\epsilon(u)-\bar{u}(u)\|<\epsilon$ for u in $\mathfrak{B}$ and $\|\bar{u}_\epsilon^{-1}(u)-\bar{u}^{-1}(u)\|<\epsilon$ for u in $\overline{\mathfrak{B}}$, where $[\bar{u}^{-1}(u), \overline{\mathfrak{B}}]$, $[\bar{u}_\epsilon^{-1}(u), \overline{\mathfrak{R}}]$ are the inverse maps of $[\bar{u}(u), \mathfrak{B}]$, $[\bar{u}_\epsilon(u), \mathfrak{R}]$, respectively. Let $\mathfrak{B}_\epsilon$ denote the correspond to $\overline{\mathfrak{B}}$ under the map $[\bar{u}_\epsilon^{-1}(u), \overline{\mathfrak{R}}]$. Then $[\bar{u}^{-1}(\bar{u}_\epsilon(u)), \mathfrak{B}_\epsilon]$ is a topological map of $\mathfrak{B}_\epsilon$ onto $\mathfrak{B}$ such that $\|\bar{u}^{-1}(\bar{u}_\epsilon(u))-u\|<\epsilon$ for u in $\mathfrak{B}_\epsilon$. Suppose that S is a surface having a representation $[\bar{x}(u), \overline{\mathfrak{B}}]$ which is AC E; consider the representation $[\bar{x}(\bar{u}_\epsilon(u)), \mathfrak{B}_\epsilon]$ for S. Denote by M the maximum of the absolute value of the ordinary Jacobian $J(u, [\bar{u}_\epsilon, \mathfrak{B}_\epsilon])$. Then a simple Jordan region $\overline{B}$ in $\overline{\mathfrak{B}}$ which is the image of a square s in $\mathfrak{B}_\epsilon$ under $[\bar{u}_\epsilon(u), \mathfrak{B}_\epsilon]$ has an area not exceeding $M\cdot|s|$. Let ${}^i\overline{T}$, ${}^iT$ denote transformations $[{}^i\bar{x}(u), \overline{\mathfrak{B}}]$, $[{}^i\bar{x}(\bar{u}_\epsilon(u)), \mathfrak{B}_\epsilon]$ respectively, for $i=1, 2, 3$. Then (see II, 1–9), since $[\bar{x}(u), \overline{\mathfrak{B}}]$ is AC E,

$$\mathcal{G}(s, {}^iT, E({}^iT, \mathfrak{B}_\epsilon^0)) = |\, {}^iT(s^0\cdot E({}^iT, \mathfrak{B}_\epsilon^0)\,| = |\, {}^i\overline{T}(\overline{B}^0\cdot E({}^i\overline{T}, \overline{\mathfrak{B}}^0)\,|$$

$$\leqq \int \mathrm{K}({}^ix, {}^i\overline{T}, \overline{B}^0)d\,{}^ix = \int_{\overline{B}^0} D(u, {}^i\overline{T})du \quad \text{for } i = 1, 2, 3.$$

Thus $[\bar{x}(\bar{u}_\epsilon(u)), \mathfrak{B}_\epsilon]$ is also AC E. Now the representation $[\bar{x}(\bar{u}(u)), \mathfrak{B}]$ need not be AC E. And if $[x(u), \mathfrak{B}]$ is an arbitrary preassigned representation for S, and ζ is any positive number, then $[\bar{u}(u), \mathfrak{B}]$ may be so chosen that $\|\bar{x}(\bar{u}(u))-x(u)\|<\zeta$ for u in $\mathfrak{B}$ (see I, 1). These results are summarized in the

Lemma. *Let S be a continuous surface possessing a representation which is* AC E. *If $[x(u), \mathfrak{B}]$ is an arbitrary representation for S, and if ϵ and ζ are any positive numbers, then there exists a simple Jordan region $\mathfrak{B}_\epsilon$, a topological map $[u_\epsilon(u), \mathfrak{B}_\epsilon]$ of $\mathfrak{B}_\epsilon$ onto $\mathfrak{B}$ such that $\|u_\epsilon(u)-u\|<\epsilon$ for u in $\mathfrak{B}_\epsilon$, and an* AC E *representation $[x_\epsilon(u), \mathfrak{B}_\epsilon]$ such that $\|x_\epsilon(u)-x(u_\epsilon(u))\|<\zeta$ for u in $\mathfrak{B}_\epsilon$.*

Choose positive numbers ϵ_n such that $\|x_0(u')-x_0(u'')\|<n^{-1}$ for any points u', u'' in $\mathfrak{B}_0$ satisfying $\|u'-u''\|<\epsilon_n$. Let B_n be a Jordan region in the interior of $\mathfrak{B}_0$ for which there exists a topological map $[u_n(u), B_n]$ of B_n onto $\mathfrak{B}_0$ such that $\|u_n(u)-u\|<\epsilon_n$ for u in B_n, $n=1, 2, \cdots$. Then $\|x_0(u_n(u))-x_0(u)\|<n^{-1}$. Thus the simple Jordan regions B_n fill up $\mathfrak{B}_0$ from the interior (see 4), the surfaces S_{0n} having representations $[x_0(u), B_n]$ satisfy $d(S_0, S_{0n})<n^{-1}$, and since clearly $eA(S_{0n})\leqq eA(S_0)$, it follows that $eA(S_{0n})$ converges to $eA(S_0)$ (see IV, 5). Now $d(S_n, S_{0n})<d(S_n, S_0)+n^{-1}$ and so the surfaces S_n admit representations $[{}_tx_n(u), B_n]$ for which $\|{}_tx_n(u)-x_0(u)\|<d(S_n, S_0)+n^{-1}$ for u in B_n, $n=1, 2, \cdots$ (see I, 1). Since the surfaces S_n have representations

which are AC E for $n=1, 2, \cdots$ (see V, **9**, 3), and since B_n is in the interior of $\mathfrak{B}_0$, it follows by the preceding lemma that there exist simple Jordan regions ${}_*\mathfrak{B}_n$ in $\mathfrak{B}_0$, topological maps $[{}_*u_n(u), {}_*\mathfrak{B}_n]$ of ${}_*\mathfrak{B}_n$ onto B_n such that $\|{}_*u_n(u) - u\| < \epsilon_n$ for u in ${}_*\mathfrak{B}_n$ and AC E representations $[{}_*x_n(u), {}_*\mathfrak{B}_n]$ such that $\|{}_*x_n(u) - x_n({}_*u_n(u))\| < n^{-1}$ for u in ${}_*\mathfrak{B}_n$. Since $\|x_0({}_*u_n(u)) - x_0(u)\| < n^{-1}$ for u in ${}_*\mathfrak{B}_n$, it follows that $\|{}_*x_n(u) - x_0(u)\| < d(S_n, S_0) + 3n^{-1}$ for u in ${}_*\mathfrak{B}_n$. The representations $[{}_*x_n(u), {}_*\mathfrak{B}_n]$ thus satisfy the hypotheses of the theorem in V, **9** and the additional assumptions in this section.

11. A proof for the theorem in V, **9** is now made, using the additional conditions, V, **10**, 1, 2. First, observe that V, **9** imply (see IV, **5**, 6) that $eV(S_0)$ is finite, and

$$\lim eA(S_n) = eA(S_0) = \int_{\mathfrak{B}_0^0} \|\mathcal{J}(u, [x_0, \mathfrak{B}_0])\| du.$$

This verifies V, **9**, 6 directly (see I, **8**). Let I be any interval in the interior of $\mathfrak{B}_0$; in view of V, **10**, 1, there exists an $n(I)$ such that I is in the interior of $\mathfrak{B}_n$ for $n > n(I)$. Define (see IV, **1**), for I in $\mathfrak{B}_0^0$.

$$^i\psi_n(I) = \int_I |\mathcal{J}(u, {}^iT_n)| du \qquad \text{for } i = 1, 2, 3;$$

$$\psi_n(I) = ({}^1\psi(I), {}^2\psi(I), {}^3\psi(I)), \qquad \Psi_n(I) = \|\psi_n(I)\| \qquad \text{for } n = 0, n > n(I).$$

From V, **9**, 3, it is seen that (see II, **9**; 1, IV, **3**)

$$^i\psi_n(I) = \int \mathrm{K}({}^ix, {}^iT_n, I^0) d\,{}^ix \quad \text{for } i = 1, 2, 3, n > n(I).$$

Since $eA(S_0)$ is finite, it follows (see IV, **4**) that $\mathrm{K}({}^ix, {}^iT_0, I^0)$ is summable, and (see V, **10**, 2; II, **2**, 2; II, **9**, 1)

$$\liminf \int \mathrm{K}({}^ix, {}^iT_n, I^0) d\,{}^ix \geqq \int \mathrm{K}({}^ix, {}^iT_0, I^0) d\,{}^ix \geqq {}^i\psi_0(I) \qquad \text{for } i = 1, 2, 3.$$

These relations give

1. $$\liminf {}^i\psi_n(I) \geqq {}^i\psi_0(I) \qquad \text{for } i = 1, 2, 3, I \text{ in } \mathfrak{B}_0^0.$$

By a known result (see R² [2, chap. II, §10]), it follows that for I in $\mathfrak{B}_0^0$,

2. $$u(I; [{}^i\psi_n, I]) = {}^i\psi_n(I) \qquad \text{for } i = 1, 2, 3,$$
$$u(I; [\Psi_n, I]) = \int_I \|\mathcal{J}(u, [x_n, \mathfrak{B}_n])\| du \qquad \text{for } n = 0, n > n(I).$$

A direct reasoning using relation 1 shows that

3. $$\liminf u(I; [\Psi_n, I]) \geqq u(I; [\Psi_0, I]) \qquad \text{for } I \text{ in } \mathfrak{B}_0^0.$$

For $n=0$, $n>n(I)$, the interval I lies in the interior of $\mathfrak{B}_0$; extend the sides of I until they meet the boundary of $\mathfrak{B}_n$ thus dividing $\mathfrak{B}_n$ into nine simple Jordan regions $I\equiv {}_0B_n, {}_1B_n, \cdots, {}_8B_n$. Clearly the ${}_hB_n$ fill up ${}_hB_0$ from the interior for $h=0, \cdots, 8$ (see IV, **10**, 1). Denote by ${}_hS_n$ the surface having the representation $[x_n(u), {}_hB_n]$ for $h=0, \cdots, 8$, $n=0$, $n>n(I)$. Since the representations $[x_n(u), {}_hB_n]$ for $n>n(I)$ are AC E and the essential areas $eA({}_hS_0)$ are finite for $h=0, \cdots, 8$, it follows that (see IV, **6**)

4.
$$eA({}_hS_n) = \int_{{}_hB_n} \|\mathcal{J}(u, [x_n, \mathfrak{B}_n])\| du \qquad \text{for } n > n(I);$$
$$eA({}_hS_0) \geqq \int_{{}_hB_n^0} \|\mathcal{J}(u, [x_0, \mathfrak{B}_0])\| du \qquad \text{for } h = 0, \cdots, 8.$$

From V, **10**, 2 it follows that ${}_hS_n$ converges to ${}_hS_0$ for $h=0, \cdots, 8$, so that (see IV, **5**)

$$\liminf eA({}_hS_n) \geqq eA({}_hS_0) \qquad \text{for } h = 0, \cdots, 8.$$

Since straight line segments form the subdivision of $\mathfrak{B}_n$ just introduced, one obtains

$$eA(S_n) = \int_{\mathfrak{B}_n^0} \|\mathcal{J}(u, [x_n, \mathfrak{B}_n])\| du = \sum_{h=0}^{8} eA({}_hS_n) \qquad \text{for } n > n(I).$$

Using the preceding relations and V, **9**, 4, one finds

$$\limsup eA({}_0S_n) = \limsup \left[eA(S_n) - \sum_{h=1}^{8} eA({}_hS_n) \right] \leqq \int_I \|\mathcal{J}(u, [x_0, \mathfrak{B}_0])\| du.$$

In view of relations 2, 3, 4, this implies that

5.
$$\lim u(I; [\Psi_n, I]) = u(I; [\Psi_0, I]) \qquad \text{for } I \text{ in } \overset{0}{\mathfrak{B}_0}.$$

Thus, if $\mathfrak{J}$ be any fixed interval in $\mathfrak{B}_0^0$ and $\mathfrak{F}$ be the class of all intervals I in $\mathfrak{J}$, it is clear from relations 1, 5 that the $[\psi_n, \mathfrak{F}]$ for $n=0$, $n \geqq n(\mathfrak{J})$ satisfy the hypotheses of the general lemma in V, **5**. In view of relation 2, therefore,

$$\lim \int_I |\mathcal{J}(u, {}^iT_n)| du = \int_I |\mathcal{J}(u, {}^iT_0)| du \qquad \text{for } I \text{ in } \overset{0}{\mathfrak{B}_0},\ i = 1, 2, 3.$$

This relation, together with the conditions in V, **10**, makes it clear that the hypotheses of the modified closure theorem in II, **14** are fulfilled, for the three sequences of transformations ${}^iT_n$: $[{}^ix_n(u), \mathfrak{B}_n^0]$, $n=0, 1, 2, \cdots$, where $i=1, 2, 3$. So ${}^iT_0$ is AC $E({}^iT_0, \mathfrak{B}_0^0)$ for $i=1, 2, 3$—that is, $[x_0(u), \mathfrak{B}_0]$ is AC E (see IV, **1**), and the theorem is established.

12. The theorem just proved permits the following addition to the results in IV, **6–9**.

THEOREM. *Assume that for a continuous surface S_0 there exists a sequence of surfaces S_n such that S_n converges to S_0, $eA(S_n)$ converges to $eA(S_0)$ which is finite, and each S_n for $n=1, 2, \cdots$ has an* AC E *representation. Then a necessary and sufficient condition that a representation $[x_0(u), \mathfrak{B}_0]$ for S_0 be absolutely continuous $(eA, \mathcal{J})$ is that $[x_0(u), \mathfrak{B}_0]$ be* AC E.

Proof. That the condition is sufficient is already established (see IV, **9**). So suppose that $[x_0(u), \mathfrak{B}_0]$ is absolutely continuous $(eA, \mathcal{J})$; then (see I, **8**)

$$eA(S_0) = \int_{\mathfrak{B}_0^0} \|\mathcal{J}(u, [x_0, \mathfrak{B}_0])\| du.$$

The S_n for $n=0, 1, 2, \cdots$ thus satisfy the hypotheses of the theorem in V, **9**, whence it follows that $[x_0(u), \mathfrak{B}_0]$ is AC E.

13. If in the theorem in V, **9**, the generalized Jacobians are replaced by the ordinary Jacobians, there results the

THEOREM. *Let S_n, $n=0, 1, 2, \cdots$ be a sequence of continuous surfaces satisfying the following conditions*:

1. *the surfaces S_n converge to S_0 (see* I, **1**);
2. *the surface S_0 has a representation $[x_0(u), \mathfrak{B}_0]$ for which the triple $J(u, [x_0, \mathfrak{B}_0])$ of ordinary Jacobians exists almost everywhere in $\mathfrak{B}_0^0$, and is summable on $\mathfrak{B}_0^0$ (see* IV, **1**);
3. *the surfaces S_n for $n=1, 2, \cdots$ have representations which are* AC E;
4. *the essential areas $eA(S_n)$ converge to $\int_{\mathfrak{B}_0^0}\|J(u, [x_0, \mathfrak{B}_0])\|du$.*

Then

5. *the representation $[x_0(u), \mathfrak{B}_0]$ is* AC E;
6. *the representation $[x_0(u), \mathfrak{B}_0]$ is absolutely continuous (eA, J)*;
7. *the essential variations $eV({}^iS_n)$ for the projection surfaces ${}^iS_n$ converge to $eV({}^iS_0)$ for $i=1, 2, 3$.*

A proof may be made by paralleling the proof for the theorem in V, **9**, using the modified closure theorem in II, **15**.

14. The preceding theorem permits the following addition to the results in IV, **10–13**. A proof is similar to that in V, **12**.

THEOREM. *Assume that for a continuous surface S_0 there exists a sequence of surfaces S_n such that S_n converges to S_0, $eA(S_n)$ converges to $eA(S_0)$ which is finite, and each S_n for $n=1, 2, \cdots$ has an* AC E *representation. Then a necessary and sufficient condition that a representation $[x_0(u), \mathfrak{B}_0]$ for S_0, for which the triple $S(u, [x_0, \mathfrak{B}_0])$ of ordinary Jacobians exists almost everywhere in $\mathfrak{B}_0^0$, be absolutely continuous (eA, J) is that $[x_0(u), \mathfrak{B}_0]$ be* AC E.

15. COROLLARY. *A necessary condition that a representation $[x(u), \mathfrak{B}]$ for a continuous surface S be absolutely continuous (A, J), where $A(S)$ is the Lebesgue area of S and $J(u, [x, \mathfrak{B}])$ is the triple of ordinary Jacobians is that*

$[x(u), \mathfrak{B}]$ be AC *E; if $[x(u), \mathfrak{B}]$ is absolutely continuous (A, J), then the Lebesgue area $A(S)$ equals the essential area $eA(S)$. A sufficient condition that a representation $[x(u), \mathfrak{B}]$ for S be absolutely continuous (A, J) is that $[x(u), \mathfrak{B}]$ be* AC *E and $eA(S)=A(S)$.*

Proof. The second assertion in this corollary has been established in IV, **14**. According to I, **6**, 2, there exists a sequence of polyhedra P_n which converge to S and for which $A(P_n)$ converges to $A(S)$. Now (see IV, **14**) each of the P_n has an AC E representation, and $eA(P_n)=A(P_n)$ for $n=1, 2, \cdots$. The remainder of the corollary now follows at once from the theorem in V, **14**.

The results of Rado and Reichelderfer cited in V, **2** are seen to be a special case of this corollary and of the theorem in V, **9** (see II, **15**; III, **4**, 2; V, **7**).

Bibliography

P. Franklin and N. Wiener

1. *Analytic approximations to topological transformations*, Trans. Amer. Math. Soc. vol. 28 (1926) pp. 762–785.

Z. de Geöcze

1. *Quadrature des surfaces courbes*, Mathematische und naturwissenschaftliche Berichte aus Ungarn vol. 26 (1908) pp. 1–88.

G. H. Hardy, J. E. Littlewood, G. Polya

1. *Inequalities*, Cambridge, 1934.

B. von Kerékjártó

1. *Vorlesungen über Topologie*, vol. 1, Berlin, 1923.

C. Kuratowski

1. *Topologie*. I, Warsaw-Lwów, 1933.

T. Rado

1. *Sur l'aire des surfaces courbes*, Acta Univ. Szeged vol. 3 (1927) pp. 131–169.
2. *Über das Flächenmass rektifizierbarer Flächen*, Math. Ann. vol. 100 (1928), pp. 445–479.
3. *On the problem of Plateau*, Ergebnisse der Mathematik und ihrer Grenzgebiete, Berlin, 1933.
4. *A lemma on the topological index*, Fund. Math. vol. 27 (1936) pp. 212–225.

T. Rado and P. Reichelderfer

1. *A theory of absolutely continuous transformations in the plane*, Trans. Amer. Math. Soc. vol. 49 (1941) pp. 258–307.
2. *On convergence in length and convergence in variation*, accepted for publication by the Duke Math. J.

S. Saks

1. *Theory of the integral*, Warsaw-Lwów, 1937.

L. Tonelli

1. *Sulla quadratura delle superficie*, Rendiconti delle sedute della reale accademia nazionale dei Lincei, Classe di Scienze fisiche, matematiche e naturali, vol. 3 (1926) pp. 357–362; pp. 445–450; pp. 633–638.

University of Chicago,
Chicago, Ill.

A THEORY FOR ORDINARY DIFFERENTIAL BOUNDARY PROBLEMS OF THE SECOND ORDER AND OF THE HIGHLY IRREGULAR TYPE

BY

RUDOLPH E. LANGER

1. **Introduction.** The boundary problems with which this discussion is concerned may be given in either the form

$$
\begin{aligned}
&u''(x) + \{p_{11}(x)\lambda + p_{10}(x)\}u'(x) \\
&\qquad\qquad + \{p_{22}(x)\lambda^2 + p_{21}(x)\lambda + p_{20}(x)\}u(x) = 0, \\
&h_{i1}(\lambda)u'(a) + h_{i2}(\lambda)u(a) + h_{i3}(\lambda)u'(b) + h_{i4}(\lambda)u(b) = 0, \qquad i = 1, 2,
\end{aligned} \tag{1.1}
$$

in which the coefficient functions $p_{ik}(x)$ and the solution $u(x)$ are scalars, or in the form

$$
\begin{aligned}
&\mathfrak{u}'(x) = \{\lambda\mathfrak{P}_1(x) + \mathfrak{P}_0(x)\}\mathfrak{u}(x), \\
&\mathfrak{H}_a(\lambda)\mathfrak{u}(a) + \mathfrak{H}_b(\lambda)\mathfrak{u}(b) = \mathfrak{o},
\end{aligned} \tag{1.2}
$$

in which the German capital letters designate matrices of the order two and the solution $\mathfrak{u}(x)$ is a vector, that is, a matrix of two rows and one column. In either form the parameter λ is to be taken as complex and unbounded, while the variable x is to be taken as real and on the finite interval (a, b). On this interval the coefficients $p_{ik}(x)$, or the elements of the matrices $\mathfrak{P}_k(x)$, are assumed to be differentiable, and such that the functions $r(x)$ which, in the case of the system (1.1), satisfy the equation

$$r^2(x) + p_{11}(x)r(x) + p_{22}(x) = 0, \tag{1.3}$$

or, in the case of the system (1.2) make the matrix

$$\{\mathfrak{P}_1(x) - r(x)\mathfrak{I}\} \tag{1.4}$$

singular, fulfill conditions to be stated below in §2. The coefficients $h_{ik}(\lambda)$ of the boundary relations in (1.1), or the elements of the matrices $\mathfrak{H}_a(\lambda)$ and $\mathfrak{H}_b(\lambda)$ in (1.2), as the case may be, are to be polynomials in λ of any degree, and may, of course, in particular be constants.

Any boundary problem of this type is compatible either for all values of λ or for no such value, or for a certain set of characteristic values which is finite or denumerably infinite. This discussion is primarily concerned with the latter case. With an infinite set of characteristic values, there exists, then,

---

Presented to the Society, September 2, 1941 under the title *A theory for ordinary linear differential boundary problems of highly irregular type*; received by the editors July 13, 1942.

an associated set of characteristic solutions, and by familiar procedures an infinite series of these solutions may be associated with an "arbitrary" function or vector. The series is then designated as a formal expansion of that function or vector, and the latter is in turn designated as the generating element of the expansion.

Contingent upon the fulfillment of certain more or less general conditions by the generating element, the behavior of an expansion, that is, its divergence, convergence, summability, value, and so on, is essentially determined by the boundary problem itself, specifically by the character of the adjustment which maintains between the differential equation and the boundary relations. This adjustment has, therefore, been made the basis for a classification of boundary problems into categories which are identified by the designations "regular," "mildly irregular" and "highly irregular."

For boundary problems of the regular type a relatively complete and familiar theory exists. The formal expansions include the classical Fourier's series as special cases, and have, broadly speaking, the salient properties of these series. Thus, in particular, they converge to the value of the generating element in an appropriately conventional sense whenever this element is integrable over the fundamental interval, and is of bounded variation in some neighborhood of the point under consideration.

Though less has been written upon boundary problems of the mildly irregular type(1), the state of their theory is roughly comparable with that of the theory of regular problems. In general the formal expansions are divergent, but are summable by means of familiar type to the values of the generating elements.

By contrast with this, nothing that may properly be referred to as a general theory has heretofore been given for boundary problems of the highly irregular type. In these the classical methods apparently lead into insurmountable difficulties, and simple examples show that these difficulties are not due to the methods alone. The literature on problems of this type is, therefore, scant. Only problems which are markedly specialized and symmetrical have been analyzed at all(2), and in them, even in the face of their specialization, the results obtained are only in slight measure comparable with those of the theory of regular problems. Existing discussions, far from

(1) Cf. M. H. Stone, *Irregular differential systems of order two and the related expansion problems*. Trans. Amer. Math. Soc. vol. 29 (1927) pp. 23–53. R. E. Langer, *The expansion problem in the theory of ordinary linear differential systems*, Trans. Amer. Math. Soc. vol. 31 (1929) pp. 868–906.

(2) Cf. J. W. Hopkins, Trans. Amer. Math. Soc. vol. 20 (1919) pp. 245–259. L. E. Ward, Trans. Amer. Math. Soc. vol. 29 (1927) pp. 716–745, ibid. vol. 32 (1930) pp. 544–557, ibid. vol. 34 (1933) pp. 417–434, Ann. of Math. (2) vol. 26 (1925) pp. 21–36, and Amer. J. Math. vol. 57 (1935) pp. 345–362. In all of these the differential equation of the problem is of a form included in $d^nu/dx^n+\{\lambda^n+\phi(x)\}u=0$, $n\geqq 3$. Also, J. I. Vass, Duke Math. J. vol. 2 (1936) pp. 151–165, in which the differential equation is $d^2u/dx^2-2\lambda\cos(p\pi/q)\cdot du/dx+\lambda^2u=0$.

applying to such formal expansions as have arbitrary generating elements, have been restricted to cases in which these elements are analytic as functions of the complex variable, and beyond that are of certain distinctive and extremely special structures.

The present paper is based upon a wholly different mode of approach to the problem. Its method is, in brief, the imbedding of the highly irregular problem in a continuous family of boundary problems of which all other members are regular. The given problem is thus approached through limiting considerations applied to existing theory. It is found on this basis that a sub-classification of the highly irregular problems into two virtual sub-categories is requisite. For problems of the first sub-category, and this includes all problems of the second order and highly irregular type that have been discussed heretofore at all, a theory is derived which is in many respects closely aligned with the existing theories for regular and mildly irregular problems. Though the expansions are non-convergent, they are shown to be summable, in certain specifically defined senses, to the values of the generating elements, whenever these latter fulfill conditions such as are familiarly imposed in the theory of Fourier's series. For problems of the second sub-category no results are derived, and it seems improbable that any expansion properties as conventionally understood inhere in problems of this type.

The discussion has been restricted to boundary problems of the second order. The motive for this, however, is to be sought only in the desire to keep the paper within its present bounds. The method set forth is evidently more generally applicable.

Chapter 1

The given boundary problem

2. **The normalization of the differential equation.** The forms of the boundary problems (1.1) and (1.2) remain unchanged under any integral linear change of the independent variable x. Since a suitable change of this kind reduces the interval (a, b) to the interval $(0, 1)$, it may be assumed without loss of generality, that $a=0$ and $b=1$. In the following this will be done. To obviate the incidence of complications which are not germane to the matter essentially at issue, it will be assumed that on this interval the coefficients $p_{kl}(x)$, $k=1, 2$; $l=0, 1, 2$; if the boundary problem is given in the form (1.1), or the elements of the matrices $\mathfrak{P}_1(x)$ and $\mathfrak{P}_0(x)$, if the problem is given in the form (1.2), are differentiable to any desired order, or at least to such orders as may be effectively called for.

If the boundary problem as given is in the form (1.1), let $r_1(x)$ and $r_2(x)$ designate the roots of the equation (1.3), and let it be supposed that these roots remain distinct on the interval $(0, 1)$. The equations

$$\frac{q_1'}{q_1} = -p_{10} - 2\frac{\phi'}{\phi},$$

$$r_1\frac{q_1'}{q_1} - r_1' = p_{21} + p_{11}\frac{\phi'}{\phi},$$

then define the functions q_1'/q_1 and ϕ'/ϕ, and if q_2 is determined from the formula

$$q_1q_2 = -\left(p_{20} + p_{10}\frac{\phi'}{\phi} + \frac{\phi''}{\phi}\right),$$

it is found that the equation obtained from the system

$$\begin{aligned} y_1'(x) &= \lambda r_1(x)y_1(x) + q_1(x)y_2(x), \\ y_2'(x) &= q_2(x)y_1(x) + \lambda r_2(x)y_2(x), \end{aligned} \tag{2.1}$$

by the elimination of the function $y_2(x)$, is identical with that obtained from the differential equation in (1.1) by making the substitution $u(x)=\phi(x)y_1(x)$. This substitution, together with the first of the equations (2.1), reduces the boundary relations of (1.1) to the forms

$$v_{i1}(\lambda)y_1(0) + v_{i2}(\lambda)y_2(0) + v_{i3}(\lambda)y_1(1) + v_{i4}(\lambda)y_2(1) = 0, \quad i = 1, 2, \tag{2.2}$$

in which the coefficients $v_{il}(\lambda)$ are again polynomials in λ. The differential system (1.1) is thus reducible to the form (2.1), (2.2). This latter may be conveniently written in matrix form, thus

$$\begin{aligned} \mathfrak{y}'(x) &= \{\lambda\mathfrak{R}(x) + \mathfrak{Q}(x)\}\mathfrak{y}(x), \\ \mathfrak{B}^{(0)}(\lambda)\mathfrak{y}(0) &+ \mathfrak{B}^{(1)}(\lambda)\mathfrak{y}(1) = \mathfrak{o}, \end{aligned} \tag{2.3}$$

in which

$$\begin{aligned} &\mathfrak{y}(x) \equiv \begin{pmatrix} y_1(x) \\ y_2(x) \end{pmatrix}, \quad \mathfrak{R}(x) \equiv \begin{pmatrix} r_1(x) & 0 \\ 0 & r_2(x) \end{pmatrix}, \quad \mathfrak{Q}(x) = \begin{pmatrix} 0 & q_1(x) \\ q_2(x) & 0 \end{pmatrix}, \\ &\mathfrak{B}^{(0)}(\lambda) \equiv (v_{ij}(\lambda)), \qquad \mathfrak{B}^{(1)}(\lambda) \equiv (v_{i,j+2}(\lambda))\text{(}^3\text{)}. \end{aligned} \tag{2.4}$$

If the boundary problem as given is in the form (1.2), let $r_1(x)$ and $r_2(x)$ be the roots of the determinant equation

$$\begin{vmatrix} p_{11}^{(1)}(x) - r & p_{12}^{(1)}(x) \\ p_{21}^{(1)}(x) & p_{22}^{(1)}(x) - r \end{vmatrix} = 0,$$

in which $(p_{ij}^{(1)}(x)) \equiv \mathfrak{P}_1(x)$, and let it be supposed that these roots are distinct on the interval (0, 1). The nonsingular matrix $\mathfrak{T}(x)$ which fulfills the relation

$$\mathfrak{P}_1(x)\mathfrak{T}(x) = \mathfrak{T}(x)\mathfrak{R}(x)$$

then exists, and the substitution $\mathfrak{u}(x) = \mathfrak{T}(x)\mathfrak{w}(x)$ gives to the differential

[3] Throughout the paper German capital letters will be used to designate square matrices of order two. Lower case German letters will correspondingly be used to designate vectors of two components, and Latin letters to denote scalars.

equation in (1.2) the form

$$\mathfrak{w}'(x) = \{\lambda\mathfrak{R}(x) + \mathfrak{P}_2(x)\}\mathfrak{w}(x),$$

with

$$\mathfrak{P}_2(x) = \mathfrak{T}^{-1}(x)\{\mathfrak{P}_0(x)\mathfrak{T}(x) - \mathfrak{T}'(x)\}.$$

The further substitution

$$\mathfrak{w}(x) = (\delta_{ij}e^{\int p_{ii}^{(2)}(x)dx})\mathfrak{y}(x)\text{[4]},$$

reduces the equation to the differential equation in (2.3), and in fact the entire boundary problem (1.2) to the form of (2.3). Since the problem, whether given in the form (1.1) or in the form (1.2) is thus reducible to (2.3), the further considerations may be confined to this latter form.

It has already been assumed that the functions $r_1(x)$ and $r_2(x)$ are distinct on the interval (0, 1). Further restrictions upon these functions, which are imposed in all existing theories of boundary problems (2.3) when the independent variable is real[5], and which are now also to be imposed herewith upon the present discussion are the following:

HYPOTHESIS 1. *On the interval* (0, 1) *the functions* $r_1(x)$, $r_2(x)$, *and* $\{r_1(x)-r_2(x)\}$, *are bounded from zero, and each of them is real except possibly for a constant complex factor.*

There are essentially two types of configuration which conform to this hypothesis, namely:

Configuration 1,

$$r_j(x) \equiv \sigma_j\rho(x), \qquad j = 1, 2, \tag{2.5}$$

in which σ_1 and σ_2 are constants different from zero with a ratio which is not real, and $\rho(x)$ is a real positive function which is bounded from zero; and

Configuration 2,

$$r_j(x) \equiv \sigma\rho_j(x), \qquad j = 1, 2, \tag{2.6}$$

in which σ is a non-vanishing constant, and $\rho_1(x)$, $\rho_2(x)$, are real functions which are bounded from each other and from zero.

To this point no stipulation has been made as to the assignment of subscripts to the functions $r_1(x)$ and $r_2(x)$. It is convenient to assign the subscripts now and henceforth in such a way that in the event of the configuration 1, the value of $\{\arg r_1(x) - \arg r_2(x)\}$ lies between 0 and π. If the relations

---

(4) Through the paper $\delta_{ij}=0$ if $i\neq j$ and $\delta_{ij}=1$ if $i=j$.

(5) For the theory when x ranges over a region of the complex plane cf. R. E. Langer, *The boundary problem of an ordinary linear differential system in the complex domain*, Trans. Amer. Math. Soc. vol. 46 (1939) pp. 151–190.

$$\begin{aligned} \Gamma_j(x) &\equiv \int_0^x r_j(x)dx, \qquad j = 1, 2, \\ \Gamma_j &= \Gamma_j(1), \end{aligned} \tag{2.7}$$

are used to define their left-hand members, it follows then in the event of configuration 1, that

$$0 < \arg \Gamma_1 - \arg \Gamma_2 < \pi. \tag{2.8}$$

In the event of the formulas (2.6) the value of $\{\arg r_1(x) - \arg r_2(x)\}$ is a multiple of π. By a suitable assignment of subscripts in this case of configuration 2, one or the other of the sets of relations

$$\begin{aligned} &\text{(a)} \qquad \arg \Gamma_1 - \arg \Gamma_2 = 0, \qquad |\Gamma_1| > |\Gamma_2|, \qquad \text{or} \\ &\text{(b)} \qquad \arg \Gamma_1 - \arg \Gamma_2 = \pi, \qquad |\Gamma_1| \geqq |\Gamma_2|, \end{aligned} \tag{2.9}$$

may therefore be achieved. It will be assumed in the following that the subscripts have been so assigned.

3. **The boundary conditions.** The components of the vector boundary relation of the system (2.3) are given explicitly in the equations (2.2), and in this, as has already been noted, each coefficient $\mathfrak{v}_{il}(\lambda)$ is a polynomial in λ, which may in particular vanish. Individually these relations are, of course, not uniquely specific, since they may be replaced by any independent linear combinations of the two without any modification of the content of the conditions as a whole being thereby induced. In the vector form of the condition, as it appears in (2.3), such a replacement is accomplished by the multiplication of the relation on the left by some nonsingular matrix, and conversely any such multiplication by a nonsingular matrix is of merely formal effect.

If in either of the relations (2.2) all four coefficients $\mathfrak{v}_{il}(\lambda)$, $l = 1, 2, 3, 4$, have some factor $(\lambda - \lambda_0)$ in common, the boundary problem is compatible at λ_0. The same, but no more, follows if they have $(\lambda - \lambda_0)$ as a multiple common factor. A reduction of the multiplicity of such a factor is, therefore, a permissible formal simplification, and in proceeding it will be assumed that such simplifications have been made, so that any common factor of the coefficients $\mathfrak{v}_{il}(\lambda)$, with $i = 1$ or $i = 2$, is a simple factor of at least one of them.

If the two relations (2.2) are linearly dependent identically in λ, the boundary problem is permanently compatible, and, from the point of view of this discussion, is without interest. That case is, therefore, to be excluded by the assumption that of the matrices

$$\mathfrak{B}_1^{(h,l)}(\lambda) \equiv \begin{pmatrix} \mathfrak{v}_{1h}(\lambda) & \mathfrak{v}_{1l}(\lambda) \\ \mathfrak{v}_{2h}(\lambda) & \mathfrak{v}_{2l}(\lambda) \end{pmatrix}, \qquad h, l = 1, 2, 3, 4; \ h \neq l, \tag{3.1}$$

at least one is not identically singular.

Let the maximum degree of the polynomials $\mathfrak{v}_{il}(\lambda)$ be designated by r.

The matrices (3.1) are, then, all expressible in the polynomial form

$$\underset{1}{\mathfrak{B}}^{(h,l)}(\lambda) = \mathfrak{C}^{(h,l,0)} + \lambda\mathfrak{C}^{(h,l,1)} + \cdots + \lambda^{\tau}\mathfrak{C}^{(h,l,\tau)},$$

with each symbol $\mathfrak{C}$ standing for a constant matrix, and with

$$\mathfrak{C}^{(h,l,\tau)} \neq \mathfrak{O},$$

for some indices (h, l). If the matrices $\mathfrak{C}^{(h,l,\tau)}$ are not all singular, let $\tau_2=\tau$. Otherwise, let a set of constant elements s_{ij} be determined such that the matrix (s_{ij}) is nonsingular, whereas

(i) $$(s_{11}, s_{12})\mathfrak{C}^{(h,l,\tau)} \neq 0,$$

for some (h, l),

(ii) $$(s_{21}, s_{22})\mathfrak{C}^{(h,l,\tau)} = 0,$$

for all (h, l). Then let τ_2 be defined as the least integer for which

$$(s_{21}, s_{22})\left\{\lambda^{\tau_2+1}\mathfrak{C}^{(h,l,\tau_2+1)} + \cdots + \lambda^{\tau}\mathfrak{C}^{(h,l,\tau)})\right\} \equiv 0,$$

for all (h, l)(6). Because of the assumption made above relative to the matrices (3.1), $\tau_2 \geqq 0$. If the boundary relation of the problem (2.3) is, then, multiplied by the matrix (s_{ij}), and if thereafter the matrices $(s_{ij})\mathfrak{B}^{(h)}(\lambda)$, $h=0, 1$, are again denoted simply by $\mathfrak{B}^{(h)}(\lambda)$, it follows that each element of these matrices is a polynomial in λ, and that when $\tau_1=\tau$, then τ_i is the maximum degree of the elements in an ith row.

It may be noted now that either one of the integers τ_1 and τ_2 may be increased by unity by the multiplication of the boundary relation on the left by the respective matrix

$$\begin{pmatrix} \lambda - \lambda_1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 0 & \lambda - \lambda_2 \end{pmatrix},$$

in which λ_1 and λ_2 are any values of the parameter for which the boundary problem is initially incompatible. It is thus a matter of an adjustment of the boundary problem to assure the relations

(3.2) $$\tau_i \geqq 1, \qquad i = 1, 2.$$

It must be noted, however, that this adjustment (which plays no role except in §28) is not wholly formal, for since the matrix factors used in achieving it are singular for a value of λ, that value is introduced as a characteristic value by the adjustment.

Finally, in virtue of the structure of the matrices $\mathfrak{B}^{(0)}(\lambda)$ and $\mathfrak{B}^{(1)}(\lambda)$ as

(6) For the purposes of multiplication vectors are always to be regarded as matrices, of one row and two columns if they are left-hand factors, and of two rows and one column if they are in the role of right-hand factors.

now arranged, it will be clear that the matrices $\mathfrak{W}^{(h)}(\lambda)$, $h=0, 1$, as defined by the formulas

$$(3.3)\qquad \begin{aligned} \mathfrak{W}^{(0)}(\lambda) &\equiv (\delta_{ij}\lambda^{-r_i})\mathfrak{B}^{(0)}(\lambda),\\ \mathfrak{W}^{(1)}(\lambda) &\equiv (\delta_{ij}\lambda^{-r_i})\mathfrak{B}^{(1)}(\lambda), \end{aligned}$$

have elements which are polynomials in $(1/\lambda)$, at least one element in a first row, and at least one element of a second row having a constant term which is not zero.

4. **The solutions of the differential equation.** Under the Hypothesis 1 of §2, the forms and structural properties of the solutions of the matrix differential equation

$$(4.1)\qquad \mathfrak{Y}'(x, \lambda) = \{\lambda\mathfrak{R}(x) + \mathfrak{Q}(x)\}\mathfrak{Y}(x, \lambda),$$

and hence of the vector differential equation of the boundary problem (2.3) may be regarded as known(7), especially insofar as large values of the parameter are concerned. Certain of these properties are relevant to the discussion at hand and may be cited as follows.

(i) With the matrix $\mathfrak{E}(x, \lambda)$ defined by the formula

$$(4.2)\qquad \mathfrak{E}(x, \lambda) \equiv (\delta_{ij}e^{\lambda\Gamma_i(x)}),$$

and with $\mathfrak{P}^{(0)}(x)$ defined to be identically the unit matrix, explicit formal procedures may be applied to determine successively the matrices of a sequence $\mathfrak{P}^{(h)}(x)$, $h=1, 2, 3, \cdots$, so that the expression

$$(4.3)\qquad \left\{\sum_{h=0}^{\infty}\lambda^{-h}\mathfrak{P}^{(h)}(x)\right\}\mathfrak{E}(x, \lambda),$$

formally satisfies the equation (4.1), namely so that upon substitution of the expression (4.3) in the place of $\mathfrak{Y}(x, \lambda)$ in the equation (4.1), the coefficients of like powers of λ in the two resulting members of the relation are in every case equal.

(ii) The infinite series (4.3) is in general divergent. However, to each λ half-plane of the set defined by the relations

$$(4.4)\qquad (n - 1/2)\pi - \arg\{\Gamma_1 - \Gamma_2\} \leqq \arg\lambda \leqq (n + 1/2)\pi - \arg\{\Gamma_1 - \Gamma_2\},$$

for integral values of n, there corresponds an actual analytic solution of the equation (4.1) which is asymptotically represented by the expression (4.3) for the values of λ in that half-plane.

(iii) In terms of any analytic nonsingular solution $\mathfrak{Y}(x, \lambda)$ of the equation (4.1) the general solution of that equation, and the general solution of the

(7) Cf. G. D. Birkhoff, and R. E. Langer, *The boundary problems and developments associated with a system of ordinary linear differential equations of the first order*. Proceedings of the American Academy of Arts and Sciences vol. 58 (1923) pp. 51–128.

vector differential equation of the boundary problem (2.3), are given, respectively, by the formulas

$$(4.5)\qquad \mathfrak{Y}(x,\lambda)\mathfrak{C}^{(1)},\qquad \mathfrak{Y}(x,\lambda)\mathfrak{c}^{(1)},$$

in which $\mathfrak{C}^{(1)}$ and $\mathfrak{c}^{(1)}$ are an arbitrary matrix and an arbitrary vector that are independent of x.

If in the first of the expressions (4.5) the matrix $\mathfrak{C}^{(1)}$ is written as $\mathfrak{Y}^{-1}(0,\lambda)\mathfrak{C}$, the general solution of the equation (4.1) is expressed in the form

$$(4.6)\qquad \mathfrak{Y}(x,\lambda)\mathfrak{Y}^{-1}(0,\lambda)\mathfrak{C}.$$

In this form the solution involved is wholly determined by the matrix $\mathfrak{C}$, since the form (4.6) is invariant under the substitution of any one nonsingular solution $\mathfrak{Y}(x,\lambda)$ for any other one. The general solution of the vector equation (2.3) may be similarly given by the formula

$$(4.7)\qquad \mathfrak{Y}(x,\lambda)\mathfrak{Y}^{-1}(0,\lambda)\mathfrak{c}.$$

Any specific one of the solutions $\mathfrak{Y}(x,\lambda)$ to which the statement (ii) above refers, defines through the formula

$$(4.8)\qquad \mathfrak{Y}(x,\lambda)\equiv\mathfrak{P}(x,\lambda)\mathfrak{E}(x,\lambda),$$

a matrix $\mathfrak{P}(x,\lambda)$ which is analytic in λ, and which, by (ii), is such that for λ in the respective half-plane of the set (4.4), the relation

$$(4.9)\qquad \mathfrak{P}(x,\lambda)\sim\mathfrak{I}+\sum_{h=1}^{\infty}\lambda^{-h}\mathfrak{P}^{(h)}(x)$$

maintains. From this it is evident that the matrix in question is nonsingular when $|\lambda|$ is sufficiently large, and that, therefore, it may be used in the role of $\mathfrak{Y}(x,\lambda)$ in the formulas (4.6) and (4.7). The asymptotic representation of the solution (4.6) or (4.7) determined by any specific matrix $\mathfrak{C}$ or vector $\mathfrak{c}$ is thus obtainable from the relations (4.8) and (4.9). This representation is valid for all large values of λ, despite the fact that the matrix $\mathfrak{P}(x,\lambda)$ to which the relation (4.9) applies is different in different half-planes (4.4), precisely by virtue of the fact that the formulas (4.6), (4.7) are invariant under replacements of the solution $\mathfrak{Y}(x,\lambda)$.

Chapter 2

The family of boundary problems

5. **The formal construction and characteristic equation of the family.** Let κ_{il}, $i=1,2;\ l=1,2,3,4$, be a set of constants, which for the instant may remain unspecified, and let ν be taken as a parameter whose range is to include the value zero. The formulas

$$(5.1)\qquad \begin{aligned}\mathfrak{B}^{(0)}(\lambda,\nu)&\equiv\mathfrak{B}^{(0)}(\lambda)+\nu(\kappa_{ij}\lambda^{r_i}),\\ \mathfrak{B}^{(1)}(\lambda,\nu)&\equiv\mathfrak{B}^{(1)}(\lambda)+\nu(\kappa_{i,j+2}\lambda^{r_i}),\end{aligned}$$

then define the matrices which appear as their left-hand members, the elements of these matrices being polynomials in λ and linear polynomials in ν. The differential system

$$(5.2)\qquad \begin{aligned} \mathfrak{y}'(x, \nu) &= \{\lambda\mathfrak{R}(x) + \mathfrak{Q}(x)\}\mathfrak{y}(x, \nu), \\ \mathfrak{B}^{(0)}(\lambda, \nu)\mathfrak{y}(0, \nu) &+ \mathfrak{B}^{(1)}(\lambda, \nu)\mathfrak{y}(1, \nu) = \mathfrak{o}, \end{aligned}$$

then defines a family of boundary problems which yields the given problem (2.3) for the parameter value $\nu=0$.

If, on the pattern of the formulas (3.3), the matrices $\mathfrak{W}^{(h)}(\lambda, \nu)$, $h=0, 1$, are now defined by the formulas

$$(5.3)\qquad \mathfrak{W}^{(h)}(\lambda, \nu) = (\delta_{ij}\lambda^{-r_i})\mathfrak{B}^{(h)}(\lambda, \nu), \qquad h = 0, 1,$$

it follows from (3.3) and (5.1) that

$$(5.4)\qquad \begin{aligned} \mathfrak{W}^{(0)}(\lambda, \nu) &\equiv \mathfrak{W}^{(0)}(\lambda) + \nu(\kappa_{ij}), \\ \mathfrak{W}^{(1)}(\lambda, \nu) &\equiv \mathfrak{W}^{(1)}(\lambda) + \nu(\kappa_{i,j+2}). \end{aligned}$$

The elements of these matrices are therefore polynomials in $(1/\lambda)$, of which the constant terms are linear in ν and all other terms are independent of ν. Moreover, at least one element in a first row and at least one element in a second row has a constant term that does not vanish when $\nu=0$.

The general solution of the differential equation of the problem (5.2) is given by the expression (4.7), a non-trivial solution being associated with a non-vanishing vector $\mathfrak{c}$. Upon substitution of this expression into the boundary relation, the latter assumes the form

$$(5.5)\qquad \mathfrak{D}(\lambda, \nu)\mathfrak{Y}^{-1}(0, \lambda)\mathfrak{c} = \mathfrak{o},$$

in which

$$(5.6)\qquad \mathfrak{D}(\lambda, \nu) \equiv \mathfrak{B}^{(0)}(\lambda, \nu)\mathfrak{Y}(0, \lambda) + \mathfrak{B}^{(1)}(\lambda, \nu)\mathfrak{Y}(1, \lambda).$$

The condition that there exist a non-vanishing vector $\mathfrak{c}$ to satisfy the equation (5.5), and hence that there exist a non-trivial solution (4.7) of the boundary problem is, therefore, evidently that the matrix (5.7) be singular, namely that

$$(5.7)\qquad D(\lambda, \nu) = 0,$$

where $D(\lambda, \nu)$ denotes the determinant of the matrix (5.6). The compatibility of the boundary problem corresponding to any specific value of ν, is thus contingent upon λ being a root of the *characteristic equation* (5.7). These roots are called the *characteristic values*.

It will be noted that the matrix $\mathfrak{D}(\lambda, \nu)$, and hence also its determinant $D(\lambda, \nu)$, depends upon the choice of the nonsingular solution $\mathfrak{Y}(x, \lambda)$ of the equation (4.1) which appears in (5.6). However, as has already been observed, any product $\mathfrak{Y}(x, \lambda)\mathfrak{Y}^{-1}(0, \lambda)$ is independent of the solution $\mathfrak{Y}(x, \lambda)$ from

which it is formed. From this it is seen at once, that the product

$$\mathfrak{D}(\lambda, \nu)\mathfrak{Y}^{-1}(0, \lambda) \tag{5.8}$$

is invariant, and that the left-hand members of the equations (5.7) formed from different solutions $\mathfrak{Y}(x, \lambda)$ differ only in their non-vanishing constant factors. The characteristic values, as roots of the equation (5.7), are thus independent of the choice of $\mathfrak{Y}(x, \lambda)$.

Since the determinant $D(\lambda, \nu)$, when formed from a solution $\mathfrak{Y}(x, \lambda)$ that is analytic in λ, is itself analytic, the number of characteristic values in any bounded portion of the complex λ-plane, and hence in particular within any circle however large, is evidently finite. For the consideration of those roots which lie outside of a suitably large circle, it is convenient to construct the equation (5.7), for λ in any half-plane of the set (4.4), from that solution to which the formulas (4.8), (4.9) apply. If the elements $a_{il}(\lambda, \nu)$, $i=1, 2$; $l=1, 2, 3, 4$, are defined, then, by the formulas

$$\begin{aligned}(a_{ij}) &\equiv \mathfrak{W}^{(0)}(\lambda, \nu)\mathfrak{P}(0, \lambda),\\ (a_{i,j+2}) &\equiv \mathfrak{W}^{(1)}(\lambda, \nu)\mathfrak{P}(1, \lambda),\end{aligned} \tag{5.9}$$

it is found that

$$\mathfrak{D}(\lambda, \nu) \equiv (\delta_{ij}\lambda^{r_i})(a_{ij} + a_{i,j+2}e^{\lambda\Gamma_j}). \tag{5.10}$$

The determinant $D(\lambda, \nu)$ is accordingly given by the formula

$$D(\lambda, \nu) \equiv \lambda^{r_1+r_2}\{A_1 - A_2e^{\lambda\Gamma_1} + A_3e^{\lambda(\Gamma_1+\Gamma_2)} - A_4e^{\lambda\Gamma_2}\}, \tag{5.11}$$

in which, if a_{i5} is interpreted as being identical with a_{i1},

$$A_l(\lambda, \nu) \equiv \begin{vmatrix} a_{1l} & a_{1,l+1} \\ a_{2l} & a_{2,l+1} \end{vmatrix}, \qquad l = 1, 2, 3, 4. \tag{5.12}$$

Since the matrices $\mathfrak{Y}(x, \lambda)$ which enter into these formulas are different for λ in different half-planes (4.4), the elements $a_{il}(\lambda, \nu)$, and the determinants $A_l(\lambda, \nu)$ are also different functions for such different values of λ. However, each solution $\mathfrak{Y}(x, \lambda)$ in question is asymptotically described as it is used by the relation (4.9), and in this relation the matrices on the right are specific and independent of λ. It follows that each element $a_{il}(\lambda, \nu)$, and likewise each determinant (5.12), may be taken as asymptotically equivalent to a respective formal power series in $1/\lambda$, and thus as subject to a single representation for all large values of λ.

It is useful, for the exploitation of certain symmetries to extend the definitions of the elements $a_{il}(\lambda, \nu)$ and of the determinants $A_l(\lambda, \nu)$ to all indices l. This may be done by the conventions

$$\begin{aligned} a_{il_1} &\equiv a_{il_2},\\ A_{l_1} &\equiv A_{l_2}, \qquad \text{for } l_1 \equiv l_2 \pmod 4).\end{aligned} \tag{5.13}$$

If the constants V_l are then defined, thus, for all integers h

$$(5.14)\qquad V_{2h-1} = (-1)^h\left\{\frac{\Gamma_2 + \Gamma_1}{2}\right\}, \qquad V_{2h} = (-1)^h\left\{\frac{\Gamma_2 - \Gamma_1}{2}\right\},$$

it is easily verifiable that the formula (5.11) may be written in the form

$$(5.15)\qquad D(\lambda, \nu) \equiv \lambda^{r_1+r_2}e^{-\lambda V_1}\sum_{l=m}^{m+3}(-1)^{l+1}A_l e^{\lambda V_l},$$

with any choice whatever of the integer m. The characteristic values other than zero, and hence all those which are numerically large, are thus roots of the equation

$$(5.16)\qquad A_m e^{\lambda V_m} - A_{m+1}e^{\lambda V_{m+1}} + A_{m+2}e^{\lambda V_{m+2}} - A_{m+3}e^{\lambda V_{m+3}} = 0.$$

6. **On the regularity or irregularity of a boundary problem.** For any given value of ν, any specific coefficient $A_l(\lambda, \nu)$ of the equation (5.16) is a function of λ. This function either vanishes identically or is asymptotically representable by a series in powers of $1/\lambda$, with a constant term that may appropriately be designated by the symbol $A_l(\infty, \nu)$. The coefficient $A_l(\lambda, \nu)$ in question will be said to be ***regular*** or ***irregular*** at the given value of ν, according as its constant term $A_l(\infty, \nu)$ is different from zero or vanishes.

If a given boundary problem is one that fulfills the hypothesis 1 in the manner of the configuration 1 of §2, the points λV_l for any four successive indices l, mark the vertices of a parallelogram centered at the origin in the complex λ-plane. The abscissas of these vertices evidently determine the magnitudes of the respective exponentials in the left-hand member of the characteristic equation (5.16), and that exponential which is associated with the vertex furthest to the right is the dominant one. Inasmuch as the orientation of the parallelogram is a function of arg λ, and any specific vertex is furthest to the right for some values of arg λ, the exponentials in the equation (5.16) may in this case all be characterized as in an obvious sense, ***potentially dominant.***

If the boundary problem fulfills the hypothesis 1 in the manner of the configuration 2, on the other hand, the points λV_l are collinear, and lie upon the segment terminated at λV_1, and λV_3 if the formulas (2.9a) apply, and on the segment terminated at λV_2 and λV_4 if the formulas (2.9b) are applicable. Since a position furthest to the right is impossible for all but the end points of the segment, only two of the exponentials which appear in the equation (5.16) are in this case potentially dominant.

The type of the boundary problem is essentially determined by those coefficients $A_l(\lambda, \nu)$ that are associated with potentially dominant exponentials in the characteristic equation. If the coefficients of the potentially dominant exponentials are all regular, the boundary problem itself is said to be of the ***regular type.*** If at least one coefficient of a potentially dominant

exponential is irregular, but no one vanishes identically, the boundary problem is said to be of the *mildly irregular* type. Finally, if in the equation (5.16) at least one coefficient of a potentially dominant exponential is identically zero, but at least two coefficients of the equation are not identically zero, the boundary problem is said to be of the *highly irregular* type.

It will be observed at once that this classification fails to account for such boundary problems as have characteristic equations with less than two non-vanishing terms. Such problems, however, have no expansion theories associated with them. For, if in the equation (5.16) just one term is non-vanishing, the number of characteristic values is clearly finite. On the other hand, if every coefficient vanishes, the equation (5.16) is evidently vacuous. The boundary problem is then compatible for all values of λ, and no characteristic values are distinguished.

The classification thus described applies in particular to the boundary problem (2.3) which was originally given, and which is identified in the family by the parameter value $\nu=0$. The discussion at hand is concerned wholly with the case in which that problem is highly irregular. Of those coefficients $A_l(\lambda, 0)$ which multiply potentially dominant exponentials in the characteristic equation, at least one is therefore to be taken as identically zero. It will be supposed, primarily for the purpose of delimiting these deductions to their present bounds, that those coefficients which do not vanish identically are regular. Although this is in fact a restrictive hypothesis, inasmuch as the case in which some non-vanishing coefficients are irregular is a more general one, the features which are engendered by such irregularities are, from the standpoint here to be maintained only secondarily germane. They constitute in the first instance the salient source of the distinctions between the regular and the mildly irregular cases.

HYPOTHESIS 2. *The given boundary problem is one for which at least one coefficient of a potentially dominant exponential in the characteristic equation vanishes identically, and for which the non-vanishing coefficients are regular and at least two in number.*

7. **Specifications upon the family of boundary problems.** Inasmuch as the constants κ_{il} introduced in §5 have remained unspecified, the boundary problem of the family associated with any value of ν different from zero has been only formally defined, and its type, in particular, has remained indeterminate. This is now to be made specific. From the formulas (4.9), (5.4) and (5.9), the evaluations

$$a_{il}(\infty, \nu) = a_{il}(\infty, 0) + \nu\,\kappa_{il}, \quad i = 1, 2;\ l = 1, 2, 3, 4,$$

are obtained. Through the relation (5.12), therefore, the expressions $A_l(\infty, \nu)$ are formally quadratic polynomials in ν in which the coefficients of the linear and quadratic terms are functions of the constants κ_{il}. It is now to be stipu-

lated that these constants be chosen so that the coefficient of each quadratic term vanishes. More precisely: *The constants* κ_{il}, $i=1, 2$; $l=1, 2, 3, 4$, *shall be such that each expression* $A_l(\infty, \nu)$ *is a linear polynomial in* ν, *and not identically zero.*

That this specification is not impossible of fulfillment in any case, may be established as follows, by the explicit display of a set of constants which have the requisite properties. Under the hypothesis 2, there exists an index p which is such that

$$A_p(\lambda, 0) \equiv 0, \qquad A_{p+1}(\infty, 0) \not\equiv 0, \tag{7.1}$$

and $A_{p+2}(\infty, 0)$, $A_{p+3}(\infty, 0)$ are not both zero. With such an index p let the constants κ_{il} be taken thus:

$$\kappa_{ip} = -a_{i,p+2}(\infty, 0), \qquad \kappa_{i,p+1} = 0, \qquad \kappa_{i,p+2} = -a_{ip}(\infty, 0), \qquad \kappa_{i,p+3} = 0,$$
$$i = 1, 2.$$

It is readily computed that with these constants

$$\begin{aligned} A_p(\infty, \nu) &= \nu A_{p+1}(\infty, 0), \\ A_{p+1}(\infty, \nu) &= A_{p+1}(\infty, 0), \\ A_{p+2}(\infty, \nu) &= A_{p+2}(\infty, 0) + \nu A_{p+3}(\infty, 0), \\ A_{p+3}(\infty, \nu) &= A_{p+3}(\infty, 0) + \nu A_{p+2}(\infty, 0), \end{aligned}$$

and since each of these expressions is a linear polynomial in ν with at least one nonzero coefficient, they evidently all have the structure prescribed.

With the coefficients $A_l(\lambda, \nu)$ thus constructed, it is evidently possible to determine in the complex ν-plane a closed neighborhood of the origin within which they are all regular except possibly at $\nu=0$. Such a neighborhood will be referred to as a *proper region* for ν, and henceforth it shall be understood that all values of ν that are brought into question lie in such a region. With the parameter so delimited the family of boundary problems is now such that each of its members associated with a value of ν different from zero is of the regular type, and only the originally given problem is irregular. In an evident sense, therefore, the given highly irregular boundary problem has been imbedded in a continuous aggregate of regular problems, and appears as analytically approachable through this aggregate by the medium of a passage of the parameter to the limiting value zero. The continuing discussion is almost exclusively concerned with considerations centering upon such an approach. Inasmuch as it is adequate to the ends sought to restrict the considerations to modes of approach in which arg ν is bounded, that restriction is to be understood henceforward.

8. Two sub-categories of highly irregular boundary problems. Rouché's theorem. The method by which a theory for boundary problems of the highly irregular type is thus to be deduced, depends essentially upon the establish-

ment of a one to one correspondence between the characteristic values and solutions of the given problem with those of the regular problems of the imbedding family, and the consequent expression of the former as limits of the latter as $\nu \to 0$. The existence of these limits as finite values is, therefore, obviously a primary requisite, and since they may or may not all exist, depending upon the individual problem at hand, a partition of the entire category of highly irregular boundary problems into sub-categories is called for. These will be distinguished by the designations A and B. Problems in which the limits in question do all exist will be allocated to the sub-category A, and to them the theory under deduction will be applicable. All highly irregular boundary problems of the second order for which any analyses at all are at present extant belong to this sub-category. On the other hand, problems in which some of the limits fail to exist will be allocated to the sub-category B. To them the theory will have no application, and it seems improbable that problems of this type admit of any expansion theory of a customary sort.

A familiar theorem([8]), upon which many of the considerations which follow are to be based, may be stated thus:

If within and on any specific closed contour of the complex λ-plane, two functions $\phi(\lambda)$ and $\psi(\lambda)$ are each analytic, and if on this contour the relation

$$|\psi(\lambda)| < |\phi(\lambda)|, \tag{8.1}$$

maintains, the equation

$$\phi(\lambda) + \psi(\lambda) = 0, \tag{8.2}$$

has precisely as many roots within the contour as has the equation

$$\phi(\lambda) = 0. \tag{8.3}$$

For future reference it will be noted here, that due to the manner in which the parameter ν enters into the structure of the functions $A_l(\lambda, \nu)$, the following may be stated.

$$\begin{aligned} &\text{If} \quad A_l(\lambda, 0) \equiv 0, \quad \text{then} \\ &\qquad A_l(\lambda, \nu) \equiv \nu\{\beta_l + \eta_l(\lambda, \nu)\}, \qquad \text{with } \beta_l \neq 0. \end{aligned} \tag{8.4}$$

$$\begin{aligned} &\text{If} \quad A_l(\lambda, 0) \not\equiv 0, \quad \text{then} \\ &\qquad A_l(\lambda, \nu) \equiv \{\alpha_l + \beta_l\nu + \eta_l(\lambda, \nu)\}, \quad \text{with } \alpha_l + \beta_l\nu \not\equiv 0. \end{aligned} \tag{8.5}$$

In either case $\eta_l(\lambda, \nu)$ designates a function that is asymptotically representable by a series in powers of $1/\lambda$ with a vanishing constant term, and otherwise with coefficients that are polynomials in ν. It is evident, therefore, that the relation

$$\lim_{\lambda \to \infty} \eta_l(\lambda, \nu) = 0, \tag{8.6}$$

([8]) Rouché's theorem. Cf. E. C. Titchmarsh, *The theory of functions*, Oxford, 1932, p. 116.

maintains uniformly as to ν. The coefficients α_l and β_l are in every case constants.

CHAPTER 3

BOUNDARY PROBLEMS OF THE SUB-CATEGORY B

9. **Problems of the configuration 2.** In the case of any boundary problem which fulfills the hypothesis 1 in the manner of the configuration 2 of §2, the ratio of the constants Γ_1 and Γ_2 is real. It will be shown that all such problems are to be allocated to the sub-category B. If the case in hand is one to which the formulas (2.9a) are applicable, the constant γ which fulfills the relation $\Gamma_1 = \gamma\Gamma_2$, is positive and greater than 1. The potentially dominant exponentials in the characteristic equation (5.16) are those in which the subscripts are odd, and at $\nu=0$ the coefficient of at least one of these is zero. Let p be chosen so that this coefficient is $A_p(\lambda, \nu)$. It is found then, that after division by the leading exponential the equation (5.16) may be written in the form

$$A_p(\lambda, \nu) - A_{p+s}(\lambda, \nu)e^{\gamma\lambda\Gamma} + A_{p+2s}(\lambda, \nu)e^{(1+\gamma)\lambda\Gamma} - A_{p+3s}(\lambda, \nu)e^{\lambda\Gamma} = 0, \tag{9.1}$$

with $s=1$, and with $\Gamma=\Gamma_2$ or $\Gamma=-\Gamma_2$ according as $p=1$, or $p=3$. If, alternatively, the problem given is one to which the formulas (2.9b) apply the value of γ which fulfills the relation $\Gamma_1 = -\gamma\Gamma_2$ is positive and at least equal to 1. In this case the potentially dominant exponentials are those with even subscripts, and if $A_p(\lambda, \nu)$ is taken as the coefficient of such a one and as vanishing at $\nu=0$, it is found that the equation (5.16) is again expressible in the form (9.1), in this instance with $s=-1$, and with $\Gamma=\Gamma_2$ or $\Gamma=-\Gamma_2$ according as $p=2$, or $p=4$. The problems of the configuration 2 may, therefore, all be analyzed by a consideration of the equation (9.1).

The case in which $\gamma=1$ may be readily disposed of. The equation (9.1) is then quadratic in $e^{\lambda\Gamma}$, and as $A_p(\lambda, \nu)$ tends to the limit zero with ν some roots $e^{\lambda\Gamma}$ and hence some characteristic values λ, become infinite. In the further considerations, in which it may now be assumed that $\gamma>1$, it is convenient to analyze separately the cases in which $A_{p+3s}(\lambda, \nu)$ does not vanish with ν, and that in which it does.

If the formulas (8.5) apply when $l=p+3s$, then since (8.4) applies when $l=p$ the characteristic equation (9.1) is expressible in the form (8.2) with

$$\phi = \beta_p - \frac{\alpha_{p+3s}}{\nu} e^{\lambda\Gamma},$$

$$\psi = \eta_p(\lambda, \nu) - (1/\nu)e^{\lambda\Gamma}\{\beta_{p+3s}\nu + \eta_{p+3s}(\lambda, \nu) + A_{p+s}(\lambda, \nu)e^{(\gamma-1)\lambda\Gamma} - A_{p+2s}(\lambda, \nu)e^{\gamma\lambda\Gamma}\}.$$

The roots of the equation (8.3) are located at the points λ_m^* given for integral values of m by the formula

$$\lambda_m^* = (1/\Gamma)\left\{2m\pi i + \log\left(\frac{\beta_p\nu}{\alpha_{p+3s}}\right)\right\}. \tag{9.2}$$

If δ is any positive constant such that $\delta|\Gamma_1| < \pi$, these roots are enclosed individually by the circles of the nonoverlapping set

$$\lambda = \lambda_m^* + \Delta\lambda, \qquad |\Delta\lambda| = \delta, \tag{9.3}$$

and it is seen at once that on any such circle

$$(1/\nu)e^{\lambda\Gamma} = (\beta_p/\alpha_{p+3s})e^{\Gamma\Delta\lambda},$$

and that as $\nu \to 0$,

$$\lambda \to \infty, \quad \text{and} \quad e^{\gamma\lambda\Gamma} \to 0.$$

Since by the first of these relations

$$\phi(\lambda) = \beta_p(1 - e^{\Gamma\Delta\lambda}),$$

on any circle (9.3), there clearly exists a positive constant M which is independent of ν, and such that for λ on the circles the relation $|\phi(\lambda)| > M$ maintains. But it is also clear from the evaluations given, that $|\psi(\lambda, \nu)| < M$ whenever $|\nu|$ is sufficiently small. For all such values of ν, therefore, the condition (8.1) is fulfilled, and it follows that each circle contains a root of the equation (8.2), namely contains a characteristic value. It is evident from (9.2), however, that each of the circles (9.3) recedes to infinity as $\nu \to 0$. The enclosed characteristic values therefore approach no finite limits.

If with $l = p + 3s$ the formulas (8.4) apply, then since they also apply with $l = p$, it follows under the hypothesis 2 that the function $A_{p+s}(\lambda, \nu)$ is given by the formula (8.5). In this case the characteristic equation (9.1) may be written in the form (8.2) with

$$\phi = \beta_p - (\alpha_{p+s}/\nu)e^{\gamma\lambda\Gamma},$$

$$\psi = \eta_p(\lambda, \nu) - (1/\nu)e^{\gamma\lambda\Gamma}\{\beta_{p+s}\nu + \eta_{p+s}(\lambda, \nu) - A_{p+2s}(\lambda, \nu)e^{\lambda\Gamma}\}$$
$$- e^{\lambda\Gamma}\{\beta_{p+3s} + \eta_{p+3s}(\lambda, \nu)\}.$$

The roots of the equation (8.3) are now located at the points

$$\lambda_m^* = (1/\gamma\Gamma)\left\{2m\pi i + \log\left(\frac{\beta_p\nu}{\alpha_{p+s}}\right)\right\},$$

and with this interpretation of λ_m^* these roots are again enclosed in the circles (9.3). On these circles it is seen that

$$(1/\nu)e^{\gamma\lambda\Gamma} = (\beta_p/\alpha_{p+s})e^{\gamma\Gamma\Delta\lambda},$$
$$\phi(\lambda) = \beta_p(1 - e^{\gamma\Gamma\Delta\lambda}),$$

and that as $\nu \to 0$

$$\lambda \to \infty, \quad \text{and} \quad e^{\lambda\Gamma} \to 0.$$

By precisely the reasoning of the previous case it is seen that each circle encloses a characteristic value and carries it to infinity as $\nu \to 0$.

The assertion that all boundary problems which conform to the configuration 2 are of the sub-category B, has thus been substantiated.

10. **Problems for which two consecutive coefficients of the characteristic equation vanish.** Since by virtue of the results of the preceding section, the continuing discussion is concerned only with boundary problems which conform to the configuration 1, the points V_l, $l=1, 2, 3, 4$, in the complex plane mark the vertices of an actual parallelogram, and each exponential in the characteristic equation is potentially dominant. Let the interior angle of this parallelogram at the vertex V_l be designated by ω_l. The cases upon which the attention is to be focused in this section, are those in which for some index l the two consecutive coefficients, $A_l(\lambda, \nu)$ and $A_{l+1}(\lambda, \nu)$ vanish with ν. Since the respective angles ω_l and ω_{l+1} are adjacent angles of the parallelogram, one of them at least does not exceed a right angle, and if this one is designated by ω_p, the index p is thereby fixed to be such that

$$A_p(\lambda, 0) \equiv 0, \qquad \omega_p \leqq \pi/2,$$

and also such that with either $s=1$ or $s=-1$, as the case may be,

$$A_{p+s}(\lambda, 0) \equiv 0.$$

With this determination of p the formulas (8.4) apply when $l=p$, $p+s$, and the formulas (8.5) do so when $l=p+2s$, $p-s$. After division by ν and by the leading exponential, the characteristic equation (5.16) is accordingly expressible in the form (8.2) with

$$\begin{aligned} \phi &= \beta_p - (\alpha_{p-s}/\nu)e^{\lambda(V_{p-s}-V_p)}, \\ \psi &= \eta_p(\lambda, \nu) - \{\beta_{p+s} + \eta_{p+s}(\lambda, \nu)\}e^{\lambda(V_{p+s}-V_p)} \\ &\quad - (1/\nu)e^{\lambda(V_{p-s}-V_p)}\{\beta_{p-s}\nu + \eta_{p-s}(\lambda, \nu) - A_{p+2s}(\lambda, \nu)e^{\lambda(V_{p+s}-V_p)}\}. \end{aligned} \tag{10.1}$$

In this instance the roots of the equation (8.3) lie at the points

$$\lambda_m^* = \frac{1}{V_{p-s} - V_p}\left\{2m\pi i + \log\left(\frac{\beta_p \nu}{\alpha_{p-s}}\right)\right\}. \tag{10.2}$$

With this interpretation of λ_m^*, these roots are enclosed in the circles of the set (9.3). On these circles

$$\frac{1}{\nu} e^{\lambda(V_{p-s}-V_p)} = \frac{\beta_p}{\alpha_{p-s}} e^{(V_{p-s}-V_p)\Delta\lambda}, \tag{10.3}$$

and hence

$$\phi = \beta_p(1 - e^{[V_{p-s}-V_p]\Delta\lambda}). \tag{10.4}$$

The quantity on the left of the relation (10.3) is thus seen to be bounded uniformly as to ν and m, and the existence of a constant M which is independent of ν and which is such that $|\phi| > M$ for all λ on the circles, is evident.

Consider now those circles of the set above which are associated with the values of m for which sm is positive. On these circles

$$
\begin{aligned}
\lambda[V_{p+s} - V_p] &= \frac{V_{p+s} - V_p}{V_{p-s} - V_p}\{2m\pi i + \log|\nu|\} \\
&\quad + [V_{p+s} - V_p]\left\{\Delta\lambda + \frac{1}{V_{p+s} - V_p}\log\left(\frac{\beta_p \nu}{\alpha_{p-s}|\nu|}\right)\right\}.
\end{aligned} \tag{10.5}
$$

With the evaluation

$$
\arg\left\{\frac{V_{p+s} - V_p}{V_{p-s} - V_p}\right\} = s\omega_p, \tag{10.6}
$$

the real part of the first term on the right of the formula (10.5) is found to be

$$
-\left|\frac{V_{p+s} - V_p}{V_{p-s} - V_p}\right|\left\{2sm\pi \sin\omega_p + \cos\omega_p \cdot \log\frac{1}{|\nu|}\right\}.
$$

This becomes negatively infinite as $\nu \to 0$, and since the remaining term on the right of the formula (10.5) is bounded, it follows that

$$
e^{\lambda[V_{p+s} - V_p]} \xrightarrow[\nu \to 0]{} 0.
$$

Since for λ on the circles in question $\lambda \to_{\nu \to 0} \infty$, the inequality $|\psi| < M$ maintains for all values of ν that are sufficiently small. For such ν, then, each of these circles contains a characteristic value, and these values become infinite as $\nu \to 0$. Any boundary problem for which two consecutive coefficients of the characteristic equation vanish must, therefore, be allocated to the subcategory B.

11. **A third type of problem of the sub-category** B. If for the given boundary problem the index p is determined so that

$$
A_p(\lambda, 0) \equiv 0, \tag{11.1}
$$

it may be assumed in this continuing discussion that

$$
A_{p-1}(\lambda, 0) \not\equiv 0, \qquad A_{p+1}(\lambda, 0) \not\equiv 0, \tag{11.2}
$$

since the alternative has been disposed of in §10. The characteristic equation (5.16), after division by ν and by the leading exponential, may, therefore, be written in the form (8.2), in which the function ϕ is as given by the formula (10.1). The corresponding function ψ is then expressible in the form

$$
\begin{aligned}
\psi &= \eta_p(\lambda, \nu) - (1/\nu)e^{\lambda(V_{p-s} - V_p)}\{\beta_{p-s}\nu + \eta_{p-s}(\lambda, \nu) \\
&\qquad - [A_{p+2s}(\lambda, \nu)e^{\lambda(V_{p-s} - V_p)} - A_{p+s}(\lambda, \nu)]e^{\lambda(V_{p+s} - V_{p-s})}\},
\end{aligned} \tag{11.3}
$$

and in these formulas the index s may be taken to be either 1 or -1. As in the preceding section, the roots of the equation (8.3) are given by the formulas (10.2), and are thus enclosed in the respective circles (9.3). Let the attention be directed upon those circles of this set that are associated with indices m for which sm exceeds a certain positive value to be further determined below, and let λ be considered upon these circles. The evaluations (10.3) and (10.4) then maintain, and for some positive constant M which is independent of ν, the function ϕ fulfills the relation $|\phi| > M$.

From the formulas (9.3) and (10.3) the equality

$$\lambda[V_{p+s} - V_{p-s}] = \left[\frac{V_{p+s} - V_p}{V_{p-s} - V_p} - 1\right]\{2m\pi i + \log|\nu|\}$$
$$+ [V_{p+s} - V_{p-s}]\left\{\Delta\lambda + \frac{1}{V_{p+s} - V_p}\log\left(\frac{\beta_p \nu}{\alpha_{p-s}|\nu|}\right)\right\},$$

may be verified. In the right-hand member of this the second term is independent of m and is bounded as to ν, whereas the first term has a real part which may be computed, with the use of (10.6), to be

$$\text{(11.4)}\quad \frac{-1}{|V_{p-s} - V_p|}\{2sm\pi|V_{p+s} - V_p|\sin\omega_p$$
$$+ [|V_{p-s} - V_p| - |V_{p+s} - V_p|\cos\omega_p]\log|\nu|\}.$$

If the constants V_l and ω_p involved in this are such that

$$\text{(11.5)}\qquad |V_{p-s} - V_p| - |V_{p+s} - V_p|\cos\omega_p > 0, \qquad s = 1, -1,$$

it is clear that the value (11.4) becomes infinite as $\nu \to 0$. The cases contrary to this are those which are here to be specifically considered.

If the boundary problem under discussion is one for which the relation (11.5) is not fulfilled, either when $s=1$ or when $s=-1$, then for such s the quantity within the brace in (11.4) is arbitrarily large when sm is sufficiently large, and the absolute value of the exponential

$$e^{\lambda(V_{p+s} - V_{p-s})}$$

is accordingly arbitrarily small uniformly in ν. It may be seen, therefore, from (11.3) and (10.3) that for all values of ν such that $|\nu|$ is suitably small and for the values of m such that sm exceeds a value appropriately large, the relation

$$|\psi| < M$$

is fulfilled. With the condition (8.1) thus met, each of the circles in question contains a characteristic value, and retains it in its interior as $\nu \to 0$. Inasmuch as the circles recede to infinity as $\nu \to 0$, it is clear that any boundary problem

which does not satisfy both of the relations (11.5) must be allocated to the sub-category B.

The failure of either one of the relations (11.5) admits of a simple geometrical interpretation. Relative to the parallelogram with vertices at the points V_l, $l=1, 2, 3, 4$, in the complex plane, the symbols ω_p, $|V_{p-1}-V_p|$, $|V_{p+1}-V_p|$, respectively designate the angle at the vertex V_p and the lengths of the adjacent sides. If ω_p is a right angle or an obtuse angle, no failure of the condition (11.5) is possible. However, if ω_p is acute a failure is possible and is articulate of the fact that one of the sides of the parallelogram adjacent to the vertex V_p is exceeded in length by the projection of the other one upon it. It will be seen at once that in such a case the diagonal $V_{p-1}V_{p+1}$ divides the parallelogram into two triangles each of which has at one of the vertices V_{p-1}, V_{p+1} an angle that is not acute. The boundary problems associated with such a configuration are, therefore, those which are allocated in this section to the sub-category B.

As has been remarked above, the theory under deduction will have no application to boundary problems assignable to the sub-category B. A hypothesis to disbar such problems from further consideration is, therefore, called for. To facilitate its enunciation, among other things, it is convenient to adopt here the relations

$$\Gamma_{2h-1} = (-1)^{h+1}\Gamma_1, \quad \Gamma_{2h} = (-1)^{h+1}\Gamma_2, \qquad h = 0, \pm 1, \pm 2, \cdots, \tag{11.6}$$

which extend the definitions of Γ_l to all indices l. It will be noted that under them

$$\Gamma_{l_1} = \Gamma_{l_2}, \quad \text{if} \quad l_1 \equiv l_2 \pmod 4,$$

and from the relations (5.14) that

$$\Gamma_l = V_{l+1} - V_l. \tag{11.7}$$

The vector Γ_l thus represents the lth side of the parallelogram with vertices at the points V_l, and for all l

$$\arg \Gamma_l - \arg \Gamma_{l+1} = \omega_l. \tag{11.8}$$

The characteristic equation may accordingly be written, with any choice of l, in the form

$$A_l(\lambda, \nu) - A_{l+1}(\lambda, \nu)e^{\lambda\Gamma_l} + A_{l+2}(\lambda, \nu)e^{\lambda(\Gamma_l+\Gamma_{l+1})} - A_{l+3}(\lambda, \nu)e^{\lambda\Gamma_{l+1}} = 0. \tag{11.9}$$

Since the values $|V_{p-1}-V_p|$, $|V_{p+1}-V_p|$, are now symbolizable by $|\Gamma_{p-1}|$, $|\Gamma_p|$, and hence have in some order the values $|\Gamma_1|$, $|\Gamma_2|$ the conditions (11.5) may evidently be expressed in the form (iii) below.

Hypothesis 3. *The given boundary problem is one whose characteristic equation, when written in the form* (11.9) *with* $\nu=0$, *fulfills the specifications*:

(i) *the ratio* Γ_1/Γ_2 *is not real*; *and*

(ii) *if* $A_p(\lambda, 0) \equiv 0$ *then* $A_{p-1}(\infty, 0) \cdot A_{p+1}(\infty, 0) \neq 0$, *and*
(iii) $|\Gamma_1| \cos \omega_p < |\Gamma_2|$, $|\Gamma_2| \cos \omega_p < |\Gamma_1|$.

Chapter 4

The characteristic values

12. The characteristic values for restricted values of ν. In the complex λ-plane the relations

$$(12.1) \qquad \begin{gathered} -\pi/2 - \arg(\Gamma_l - \Gamma_{l+1}) \leqq \arg \lambda < -\pi/2 - \arg(\Gamma_l + \Gamma_{l+1}), \\ |\lambda| \geqq N, \end{gathered}$$

define, for each index l and for any non-negative real constant N, a region which is to be denoted by $S_l(N)$. Any four consecutive regions of this set cover the part of the plane which lies outside of the circle with radius N centered at the origin. The asymptotic distribution of the characteristic values may therefore be determined by a study of their distribution in the region (12.1) with the index l unspecified.

The characteristic equation of the family of boundary problems has been written in the form (11.9). If the abbreviations

$$(12.2) \qquad c_l(\nu) \equiv A_l(\infty, \nu),$$

are adopted, the further relations

$$(12.3) \qquad A_l(\lambda, \nu) \equiv c_l(\nu)\{1 + \chi_l(\lambda, \nu)\},$$

define the functions $\chi_l(\lambda, \nu)$ here involved, and these are evidently all arbitrarily small, uniformly as to ν in any proper ν-region, when λ lies in a region (12.1) in which N is sufficiently large.

Under the relation (2.8), which maintains in all boundary problems now under consideration, the constants of the set

$$\sin \arg \left\{ \frac{\Gamma_l \pm \Gamma_{l+1}}{\Gamma_{l+1}} \right\}, \qquad l = 1, 2, 3, 4,$$

are all positive. Let a designate the smallest one of these constants, and let a_1 be any positive constant less than a. It is then easily verified that the relation

$$(12.4) \qquad |e^{\lambda \Gamma_{l+1}}| \leqq e^{-a|\lambda|}, \qquad \text{for } \lambda \text{ in } S_l(N),$$

maintains, and that when ν is restricted to a part of its proper region in which

$$(12.5) \qquad |\nu| \geqq e^{-a_1 N},$$

then the values

$$\left| \frac{1}{c_{l+j}(\nu)} e^{\lambda \Gamma_{l+1}} \right|, \qquad j = 0, 1, \tag{12.6}$$

are arbitrarily small in $S_l(N)$, if N is sufficiently large.

With the symbol r standing in the place of either 0 or 1, consider the functions $\phi(\lambda, \nu)$ and $\psi(\lambda, \nu)$ defined by the formulas

$$\phi(\lambda, \nu) \equiv 1 - \frac{c_{l+1}(\nu)}{c_l(\nu)} e^{\lambda\Gamma_l} + r\left\{\theta_0 - \frac{c_{l+1}(\nu)}{c_l(\nu)} e^{\lambda\Gamma_l}\theta_1\right\},$$

$$\psi(\lambda, \nu) \equiv (1 - 2r)\left\{\theta_0 - \frac{c_{l+1}(\nu)}{c_l(\nu)} e^{\lambda\Gamma_l}\theta_1\right\},$$

in which

$$\theta_0(\lambda, \nu) \equiv \chi_l(\lambda, \nu) - \frac{A_{l+3}(\lambda, \nu)}{c_l(\nu)} e^{\lambda\Gamma_{l+1}},$$

$$\theta_1(\lambda, \nu) \equiv \chi_{l+1}(\lambda, \nu) - \frac{A_{l+2}(\lambda, \nu)}{c_{l+1}(\nu)} e^{\lambda\Gamma_{l+1}}.$$

If λ^* is taken to designate any zero of the function $\phi(\lambda, \nu)$ in the region $S_l(N)$, the equation $\phi(\lambda^*, \nu) = 0$ may be written in the form

$$\frac{c_{l+1}(\nu)}{c_l(\nu)} e^{\lambda^*\Gamma_l} = \frac{1 + r\theta_0(\lambda^*, \nu)}{1 + r\theta_1(\lambda^*, \nu)}. \tag{12.7}$$

On the circle

$$\lambda = \lambda^* + \Delta\lambda, \qquad |\Delta\lambda| = \epsilon, \tag{12.8}$$

in which ϵ is positive and less than the smaller of the numbers $\pi/|\Gamma_j|$, $j = 1, 2$, but otherwise arbitrary, the relation (12.7) yields the evaluation

$$\frac{c_{l+1}(\nu)}{c_l(\nu)} e^{\lambda\Gamma_l} = \frac{1 + r\theta_0(\lambda^*, \nu)}{1 + r\theta_1(\lambda^*, \nu)} e^{\Gamma_l \Delta\lambda},$$

and since θ_0 and θ_1 are arbitrarily small over the region $S_l(N)$, it follows that on the circle (12.8) the function $\phi(\lambda, \nu)$ differs by arbitrarily little from the value $(1 - e^{\Gamma_l\Delta\lambda})$, while the function $\psi(\lambda, \nu)$ is arbitrarily small. Since a relation (8.1) thus maintains upon the circle (12.8), the equations (8.2) and (8.3) have the same numbers of roots within it.

Now when $r = 1$ the equation (8.3) is the characteristic equation and (8.2) is the equation

$$1 - \frac{c_{l+1}(\nu)}{c_l(\nu)} e^{\lambda\Gamma_l} = 0, \tag{12.9}$$

whereas when $r = 0$ the roles of these two equations are reversed. It follows

that the roots

$$\lambda^*_{l,m} = (1/\Gamma_l)\left\{2m\pi i - \log \frac{c_{l+1}(\nu)}{c_l(\nu)}\right\}, \tag{12.10}$$

of the equation (12.9) which lie in the region $S_l(N)$, may be set into one to one correspondence with the characteristic values in that region, with corresponding elements within a distance ϵ of each other. Since the points $\lambda^*_{l,m}$ are spaced at distances exceeding 2ϵ from each other, it must be concluded that the characteristic values in the region $S_l(N)$ are all simple, and that they are enumerable and denotable in the manner $\lambda_{l,m}$ so that

$$|\lambda_{l,m} - \lambda^*_{l,m}| < \epsilon, \qquad \text{for } \lambda_{l,m} \text{ in } S_l(N). \tag{12.11}$$

Inasmuch as the constant ϵ may be taken to be arbitrarily small, and the relation (12.11) is nevertheless fulfilled when N is sufficiently large, the use of the symbolism of asymptotic representation, namely

$$\lambda_{l,m} \sim \lambda^*_{l,m} \tag{12.12}$$

is evidently justified. The entire set of characteristic values is clearly enumerable, since those which lie within any circle of radius N centered at the origin are finite in number, while those outside such a circle stand in correspondence with the enumerable sets (12.10) with $l=1, 2, 3, 4$.

The relation (12.5) restricts the parameter ν from a neighborhood of the origin $\nu=0$. This prohibited neighborhood can, however, be made arbitrarily small by the choice of a value of N that is sufficiently large. Irrespective of how small the proper region to which ν is initially confined may be, therefore, the considerations above are applicable for a range of values of ν that is not empty when the characteristic values concerned are remote enough from the origin of the λ-plane. Now for ν in a suitably small region, a comparison of the formulas (12.2) with (8.4) and (8.5) shows that when $\alpha_l \neq 0$ then the difference $\{\log c_l(\nu) - \log \alpha_l\}$ is arbitrarily small, whereas when $\alpha_l = 0$ then $\log c_l(\nu) = \log(\beta_l \nu)$. It may be drawn from the relations (12.11), therefore, that

$$|\lambda_{l,m} - \overset{(0)}{\lambda}_{l,m}| < 2\epsilon, \qquad \text{for } \lambda_{l,m} \text{ in } S_l(N), \tag{12.13}$$

where

$$\overset{(0)}{\lambda}_{k-j,m} = (1/\Gamma_{k-j})\left\{2m\pi i - (-1)^j \log \frac{\alpha_{k+1-2j}}{\beta_k \nu}\right\}, \qquad j = 0, 1, \tag{12.14}$$

when k is any index for which $A_k(\lambda, 0) \equiv 0$, and

$$\overset{(0)}{\lambda}_{q-j,m} = (1/\Gamma_{q-j})\left\{2m\pi i - (-1)^j \log \frac{\alpha_{q+1-2j}}{\alpha_q}\right\}, \qquad j = 0, 1, \tag{12.15}$$

when q is an index for which $A_{q-1}(\lambda, 0) \neq 0$, $A_q(\lambda, 0) \neq 0$, and $A_{q+1}(\lambda, 0) \neq 0$. It will be observed that the points (12.15) are constant as to ν, and hence that any characteristic values represented by them through the relation (12.13) are asymptotically constant. Under the hypotheses made, however, at most one index q can exist, and there may be no such index at all for the boundary problem under consideration. In at least two and possibly in all four of the regions $S_l(N)$, $l=1, 2, 3, 4$, the characteristic values accordingly refer through the relations (12.13) to points of the respective sets (12.14), and so depend in an essential manner upon ν.

It will be observed for later reference that insofar as an index k is concerned to which the formulas (12.14) apply, the reasoning epitomized in the relations (12.13) would be in no way affected if the function $A_{k+2}(\lambda, \nu)$ were replaced by 0, and the functions $A_l(\lambda, \nu)$ for $l=k-1, k, k+1$, were replaced by their leading terms as those are given in the formulas (8.4) and (8.5). These replacements substitute the equation

$$\beta_k \nu - \alpha_{k+1} e^{\lambda \Gamma_k} - \alpha_{k-1} e^{\lambda \Gamma_{k+1}} = 0, \tag{12.16}$$

in the place of the characteristic equation. In the regions $S_{k-1}(N)$ and $S_k(N)$, therefore, the roots of this equation are also represented asymptotically by the points of the sets (12.14).

13. **On critical values of λ and ν.** By virtue of the hypothesis 3 the boundary problem at hand is one for which the relations (11.2) maintain if the index p is suitably determined. Let such a determination of p be fixed upon, and throughout this section let it be understood that k is used to stand at will for either p or $p+2$. For these values of k, the equations (12.16) are to be considered in the respective λ half-planes $S_{k-1,k}$, of which each consists of the pair of adjacent sectors $S_{k-1}(0)$ and $S_k(0)$.

If, for any value of ν, the equation (12.16) admits of a multiple root in the half-plane $S_{k-1,k}$, that root is a zero of the derived function

$$\left\{\alpha_{k+1}\Gamma_k e^{\lambda\Gamma_k} + \alpha_{k-1}\Gamma_{k+1} e^{\lambda\Gamma_{k+1}}\right\}. \tag{13.1}$$

It is, therefore, a point of the set

$$\lambda^{(k,m)} = \frac{1}{\Gamma_{k+1} - \Gamma_k}\left\{(2m+1)\pi i + \log \frac{\alpha_{k+1}\Gamma_k}{\alpha_{k-1}\Gamma_{k+1}}\right\}, \qquad m \geqq m_k, \tag{13.2}$$

with the integer m_k such that it lies in the region in question. Upon substitution of the values (13.2) into the equation (12.16), the respectively corresponding values of ν are found to be given by the formulas

$$\nu^{(k,m)} = H_k e^{-m\Omega}, \qquad k = p,\ p+2, \tag{13.3}$$

in which each coefficient H_k is a (complex) constant independent of m, whereas

$$\Omega = \frac{-2\pi i \Gamma_p}{\Gamma_{p+1} - \Gamma_p}. \tag{13.4}$$

The equation (12.16) obviously defines ν as a single-valued analytic function of λ. For the deductions at hand, however, the inverse relationship, namely the dependence of λ upon ν is of more immediate consequence. In the map defining this dependence the points (13.3) are branch points. According as the domain of ν includes these points or excludes them, the equation (12.16) may be regarded as defining $\lambda(\nu)$ as an infinitely many-valued function, or as defining its infinity of roots as distinct single-valued functions of ν. Of these alternatives the latter one is to be adopted, and the points (13.2) and (13.3) are to be referred to henceforth as *critical values* of λ and ν, respectively. It is to be shown, among other things, that there exist in the domain of ν paths of approach to the origin which avoid the critical values, and in fact that there exist such paths along which $|\nu|$ varies monotonically and arg ν varies within an arbitrarily prescribed positive range, and along which the roots of the equation (12.16) are uniformly bounded from the critical values of λ.

The formula (13.4), together with the hypothesis 3, assures that both the real and the pure imaginary components of the constant Ω are positive. For $k=p$ and for $k=p+2$, therefore, the points of the respective set (13.3) lie upon a logarithmic spiral which winds in upon the point $\nu=0$, the points given by successive indices m being spaced along this spiral at regular angular intervals of magnitude equal to the imaginary part of Ω. Let κ_1 be defined as the smallest positive constant of the set

$$\left|\frac{\Gamma_k}{\Gamma_q}e^{s\Omega} - 1\right|, \qquad k = p, p+2; q = p, p+2; s = 0, \pm 1, \pm 2, \cdots.$$

The relation

$$\left|\frac{\nu^{(k,m)} - \nu^{(q,r)}}{\nu^{(q,r)}}\right| \geqq \kappa_1, \tag{13.5}$$

maintains then for every pair of distinct critical values $\nu^{(k,m)}$ and $\nu^{(q,r)}$, irrespective of whether they lie upon the same or different spirals.

Together with a prescription of continuity at $z=0$, the formula

$$F(z) \equiv \frac{\Gamma_{p+1}(e^{\Gamma_p z} - 1) - \Gamma_p(e^{\Gamma_{p+1} z} - 1)}{(\Gamma_{p+1} - \Gamma_p)z^2},$$

defines $F(z)$ as a function of z which is analytic over the finite z-plane. This function is, therefore, in particular bounded in the unit circle, and M may accordingly be chosen as a constant such that $M \geqq 1$, and

$$|F(z)| < M, \qquad \text{for } |z| \leqq 1.$$

With any prescribed value $\arg \nu_0$, and with any positive constant δ that is exceeded by both of the constants π and κ_1, let $\sum_\delta$ designate the sector

$$|\arg \nu - \arg \nu_0| < \delta/2.$$

With such a value of δ chosen, the relations

$$(13.6) \qquad |\lambda - \lambda^{(k,m)}| < \left\{\frac{\sin(\delta/2)}{M}\right\}^{1/2}, \qquad k = p, p+2; m \geqq m_k,$$

define in the λ-plane a set of circular regions with fixed radii, and centered at the critical points (13.2). Through the relation (12.16) these regions are mapped upon respective neighborhoods of the points $\nu^{(k,m)}$. These will be referred to briefly as *critical neighborhoods.*

From the relations (12.16) and (13.2), it may be drawn without difficulty that

$$\frac{\nu - \nu^{(k,m)}}{\nu^{(k,m)}} = z^2F(z), \quad \text{with} \quad z = (-1)^{(k-p)/2}(\lambda - \lambda^{(k,m)}).$$

It follows from this that every value λ within a region (13.6) corresponds to a value of ν such that

$$(13.7) \qquad \left|\frac{\nu - \nu^{(k,m)}}{\nu^{(k,m)}}\right| \leqq \sin\left(\frac{\delta_1}{2}\right), \quad \text{with} \quad \delta_1 < \delta,$$

namely, that the critical neighborhood of the point $\nu^{(k,m)}$ is wholly within the respective circle of the set (13.7). Since $\delta_1 < \kappa_1$ and $\delta_1 < \pi$, it is clear on the one hand, because of the relation (13.5), that no two of the circles (13.7) have any points in common, and on the other hand, directly from the formula (13.7), that no one of them includes the point $\nu = 0$. Since each circle furthermore subtends at $\nu = 0$ the angle δ_1, which is less than the angle of the sector $\sum_\delta$, the following facts are easily verified. If from the sector $\sum_\delta$ all points which belong to any circle of the set (13.7) are deleted, the remainder of the sector is a connected region within which there exist continuous paths of approach to the vertex $\nu = 0$ along which $|\nu|$ steadily decreases. This is what was to be shown. Since along such a path ν remains in the chosen sector, the oscillation of $\arg \nu$ does not exceed the prescribed value δ, and since ν enters no circle (13.7), no root of the equation (12.16), either with $k = p$ or with $k = p+2$, enters into a region of the set (13.6). The roots of the equations (12.16) thus remain uniformly bounded from the critical λ-values. Paths in the ν-plane having the properties enumerated will be referred to henceforth as ***regular paths*** for ν.

It may be noted incidentally that the cases in which the imaginary component of the constant Ω/π is rational are peculiarly simple. From the formulas (13.3) it may be seen that the critical points $\nu^{(k,m)}$ then all lie upon a

finite number of rays from the origin of the ν-plane. The circles (13.7) including the critical neighborhoods are therefore centered upon these lines, and any section $\sum_s$ accordingly includes *rectilinear* paths of approach to $\nu=0$ that are regular.

It is familiar, and can easily be proved, that any exponential sum all of whose zeros occur at points of the set (13.2) is uniformly bounded from zero if λ is uniformly bounded from the points in question. Any quotient obtained by the division of a function (13.1) by one of the exponentials which it involves is such a sum, and is therefore bounded from zero when λ remains outside of the circles (13.6). With any regular path for ν there may, therefore, be associated a positive constant of ρ which is such that for all values of ν upon the path the relations

$$(13.8)\qquad \left| \alpha_{k+1}\Gamma_k e^{\lambda(\Gamma_k-\Gamma_{k+j})} + \alpha_{k-1}\Gamma_{k+1}e^{\lambda(\Gamma_{k+1}-\Gamma_{k+j})}\right| \geqq \rho, \qquad k = p,\ p+2;\ j = 0,\ 1,$$

are fulfilled by every root of an equation (12.16).

14. **The loci of the roots of an equation** (12.16). As the parameter ν varies along any regular path, the roots of the equation (12.16) with $k=p$, in the respective half-plane $S_{p-1,p}$, remain distinct and trace out continuous loci in the λ-plane. It is to be shown that there exists for each of these loci a finite *terminal point* corresponding to the parameter value $\nu=0$, and hence that every root of the equation in question approaches a finite limit as $\nu\to 0$.

The change of variable and parameter from λ and ν to $z(=x+iy)$ and μ, as given by the relations

$$(14.1)\qquad \begin{aligned} z &= (i/2)\left\{\lambda[\Gamma_p - \Gamma_{p+1}] + \log\frac{\alpha_{p+1}}{\alpha_{p-1}}\right\}, \\ \mu e^{i\theta} &= \beta_p\nu \exp\left\{\frac{\Gamma_{p+1}\log\alpha_{p+1} - \Gamma_p\log\alpha_{p-1}}{\Gamma_p - \Gamma_{p+1}}\right\}, \end{aligned}$$

with $\mu\geqq 0$, transforms the equation (12.16) into the equation

$$(14.2)\qquad e^{-z(\beta+i\alpha+i)-i\theta} + e^{-z(\beta+i\alpha-i)-i\theta} = \mu,$$

with

$$(14.3)\qquad \alpha - i\beta = \frac{\Gamma_p + \Gamma_{p+1}}{\Gamma_p - \Gamma_{p+1}}.$$

If z_0 indicates the point corresponding to $\lambda=0$, the region $S_{p-1,p}$ is transformed into the half-plane

$$(14.4)\qquad -\pi + \tan^{-1}\beta/\alpha \leqq \arg(z - z_0) \leqq \tan^{-1}\beta/\alpha,$$

and this includes all except possibly a finite segment of the positive axis of reals.

From the equality of the pure imaginary components of its two members, and the equality of their absolute values, the complex equation (14.2) may be made to yield the pair of real cartesian equations

$$(14.5)\quad \begin{aligned} &\text{(a)}\quad e^{y}\sin\left\{[1+\alpha]x+\beta y+\theta\right\} = e^{-y}\sin\left\{[1-\alpha]x-\beta y-\theta\right\},\\ &\text{(b)}\qquad 4e^{2\alpha y}\left\{\cos^2 x+\sinh^2 y\right\} = \mu^2 e^{2\beta x}.\end{aligned}$$

Since μ is proportional to $|\nu|$, it may be taken to fill the role of the parameter. Along any regular path, θ, which differs from arg ν by a constant, is then determined as a function of μ, the oscillation of θ being less than δ for any path in a sector $\sum_\delta$. By virtue of the hypothesis 3, the real constants α, β, which appear in the equations (14.5), and which are defined by the relation (14.3), are such that

$$-1<\alpha<1,\qquad 0<\beta.$$

It may be noted, however, that in the equations (14.5) an interchange of α and $-\alpha$ may be achieved by the substitution of $-y$ and $-\theta$ in the place of y and θ. Since any result derived for $\alpha>0$ may, therefore, be translated to apply when $\alpha<0$, there is no essential loss of generality in assuming for the explicit discussion that $\alpha\geqq 0$, and this will be done in the following.

With any choice of an initial parameter value ν_0, which is such that for the associated value θ_0 the constant

$$\frac{\theta_0+(\pi/2)\alpha}{(1-\alpha)\pi}$$

is not an integer, it is possible to associate an integer n_0 such that for all real constants δ_0 which are numerically sufficiently small, the relations

$$(14.6)\qquad n_s+\frac{|\delta_0|}{\pi}<\frac{(q_s+1/2)\pi\alpha+\theta_0+\delta_0\alpha}{(1-\alpha)\pi}<(n_s+1)-\frac{|\delta_0|}{\pi},$$

are fulfilled when $s=0$ with $q_0=0$. If the case is one in which $\alpha>0$, there exist then a pair of positive increasing sequences of integers $\{q_s\}$ and $\{n_s\}$, for which the relations (14.6) are fulfilled when $s=1, 2, 3, \cdots$. It thereupon follows further, again if $|\delta_0|$ is sufficiently small, that the relations

$$(14.7)\qquad n_s+\frac{|\delta_0|}{\pi}<\frac{(p_s+1/2)\pi\alpha+\theta_0+\delta_0\alpha}{(1+\alpha)\pi}<(n_s+1)-\frac{|\delta_0|}{\pi},$$

are fulfilled by the integers of a third increasing sequence $\{p_s\}$. Let such sequences relative to the chosen constant δ_0 be fixed upon. If the case is one in which $\alpha=0$, these sequences may be taken arbitrarily, since the relations (14.6), (14.7) imply no specifications for them.

Consider the relations

$$
(14.8)\quad
\begin{aligned}
&\text{(a)}\quad y \geqq 0,\quad (1+\alpha)x + \beta y = (p_s + 1/2)\pi - \theta_0 + \delta_0;\\
&\text{(b)}\quad y = 0,\quad x \text{ between } \frac{(p_s + 1/2)\pi - \theta_0 + \delta_0}{1+\alpha} \text{ and } \frac{(q_s + 1/2)\pi + \theta_0 + \delta_0}{1-\alpha};\\
&\text{(c)}\quad y \leqq 0,\quad (1-\alpha)x - \beta y = (q_s + 1/2)\pi + \theta_0 + \delta_0.
\end{aligned}
$$

In the (x, y)-plane the first and third of these define half-lines which terminate upon the axis of x, and the second is the segment intercepted by them upon this axis. The set of relations as a whole therefore defines a broken line which divides the region (14.4) into two parts, in the one of which, to be denoted by $Z_s(\delta_0)$, the abscissas are bounded above. It is clear that a region $Z_s(\delta_0)$ with a larger index includes any one with a smaller index, and that the bound upon the abscissas increases indefinitely with s. Let it be assumed now that the path of ν lies in a sector $\sum_s$ for which δ fulfills the conditions imposed upon $|\delta_0|$ above. It is to be shown for the equations (14.5) that every root (x, y) which initially lies within any region $Z_s(\delta)$ remains within that region as $\mu \to 0$, and that any root which is initially outside of any region $Z_s(-\delta)$ remains outside.

Consider any root in a position in which its ordinate is positive. For this position the equation (14.5a) shows that the sine function in the left-hand member of that equation is numerically less than unity, and hence that the value of $\{[1+\alpha]x+\beta y+\theta\}$ is not an odd multiple of $\pi/2$. With a suitable determination of s, therefore,

$$(14.9)\qquad (1+\alpha)x + \beta y < (p_s + 1/2)\pi - \theta_0 + \delta,$$

where p_s is a member of the sequence so designated through the relation (14.7) in association with the value of $\delta_0 = \delta$. The point (x, y), therefore, lies in the region $Z_s(\delta)$, and since the relation (14.9) maintains while $y>0$, it is clear that the root cannot issue from this region across the boundary (14.8a). Similarly with s properly redetermined and p_s a member of the respective sequence associated with the value $\delta_0 = -\delta$ through the relation (14.7) it is assured that

$$(p_s + 1/2)\pi - \theta_0 - \delta < (1+\alpha)x + \beta y.$$

The root (x, y) thus lies initially outside of the region $Z_s(-\delta)$ with this index s, and since the reasoning employed above shows it to remain outside so long as $y>0$, it is evident that no root may enter any such region across the portion (14.8a) of its boundary.

If in any of its positions the ordinate of a root (x, y) is negative, the equation (14.5a) shows that the value of $\{[1-\alpha]x-\beta y-\theta\}$ is not an odd multiple of $\pi/2$, and hence that

$$(q_{s_1} + 1/2)\pi + \theta_0 - \delta < [1 - \alpha]x - \beta y < (q_s + 1/2)\pi + \theta_0 + \delta,$$

provided s and s_1 are properly determined, and the sequences $\{q_s\}$ and $\{q_{s_1}\}$ are associated with the values $\delta_0 = \delta$ and $\delta_0 = -\delta$ respectively, through the relations (14.6). Since this configuration maintains so long as $y<0$, it follows that no root may either issue from a region $Z_s(\delta)$ or enter into a region $Z_s(-\delta)$ across a boundary (14.8c).

Finally, upon setting $y=0$ the equation (14.5a) is found to reduce to the form

$$\cos x \sin (\alpha x + \theta) = 0.$$

Of the roots of this equation those that are zeros of the factor $\cos x$ lie at points of the set

$$[(r + 1/2)\pi, 0], \qquad r = 0, 1, 2, \cdots . \tag{14.10}$$

They are shown by the equation (14.5b) to be uniquely associated with the parameter value $\mu=0$, and thus, as points of loci which are traced out as $\mu \to 0$, they are *terminal points*, and not points at which the loci actually cross the axis of x. Such crossing points must accordingly be zeros of the factor $\sin (\alpha x+\theta)$, and hence points at which the respective values of $\{\alpha x+\theta_0\}$ differ from integral multiples of π by less than the amount $|\delta_0|$. If $\alpha>0$ no such point lies on any segment (14.8b), either for $\delta_0=\delta$ or for $\delta_0=-\delta$, since the relations (14.6) and (14.7) insure that on any such segment

$$n_s\pi + |\delta_0| < \alpha x + \theta_0 < (n_s + 1)\pi - |\delta_0|.$$

On the other hand if $\alpha=0$ there exist no such points at all, as may be seen from the relation (14.6) with $s=0$. Since a root may, therefore, neither issue from a region $Z_s(\delta)$ nor enter into a region $Z_s(-\delta)$ over the boundary (14.8b), the assertion above has been substantiated.

This deduction admits of two specific and pertinent conclusions. In the first instance, since every root remains within some region $Z_s(\delta)$, its abscissa is subject to some upper bound. By the equation (14.5b), therefore, it approaches a limit as $\mu \to 0$, and this limit is a point of the set (14.10). In the second instance, since no root may enter into any region $Z_s(-\delta)$, it follows that the distance of any root from the point z_0 is subject to a lower bound, and that this bound is arbitrarily large for any root which is sufficiently distant at any specific value of μ.

As reinterpreted into terms of the variables λ and ν through the formulas (14.1), the results may be formulated thus. The roots λ' of the equation (12.16) with $k=p$ which lie in the half-plane $S_{p-1,p}$, all approach finite limits as $\nu \to 0$, and these limits are all points of the set

$$\frac{-1}{\Gamma_p - \Gamma_{p+1}}\left\{(2r + 1)\pi i + \log \frac{\alpha_{p+1}}{\alpha_{p-1}}\right\}. \tag{14.11}$$

Any such root λ' which at any specific value of ν is sufficiently large in absolute value, remains arbitrarily large as $\nu \to 0$.

15. **The characteristic values as $\nu \to 0$.** In §12 it was deduced that if there exists for the given boundary problem an index q, such that $A_l(\lambda, 0) \not\equiv 0$ for $l = q-1, q, q+1$, then the characteristic values which lie in the regions $S_{q-1}(N)$ and $S_q(N)$ are represented by the points of the sets (12.10) with $l = q-1, q$. Insofar as these deductions were concerned, the restriction (12.5) upon the parameter ν was wholly dispensable, since the relation (12.4) insures that the values (12.6) are arbitrarily small when N is suitably large, without recourse to the restriction upon ν. The representations of the characteristic values by the relations (12.13), (12.15) are, therefore, not only valid as stated in §12, but maintain as $\nu \to 0$.

The contrary is the case insofar as the characteristic values are concerned, which lie in any region $S_{p-1}(N)$ or $S_p(N)$ with an index p for which $A_p(\lambda, 0) \equiv 0$. That these values cannot be represented for unrestricted values of ν by the points (12.14) through the relation (12.13), is, in fact, immediately evident, since these points recede to infinity as $\nu \to 0$. For the deductions culminating in the relations (12.13), (12.14) the restriction (12.5) was, therefore, essential, and it accordingly remains to deduce for the characteristic values in any half-plane $S_{p-1,p}$ a representation which maintains as $\nu \to 0$.

For λ in a region $S_p(N)$, let the function $\phi(\lambda)$ be defined by the formula

$$\phi(\lambda) \equiv \beta_p \nu e^{-\lambda \Gamma_p} - \alpha_{p+1} - \alpha_{p-1} e^{\lambda(\Gamma_{p+1} - \Gamma_p)}. \tag{15.1}$$

With $\lambda = \lambda' + \Delta\lambda$, in which λ' is any root of the equation (12.16), an alternative formula for this function is

$$\phi(\lambda) = \{\phi_1(\lambda') - \phi_2(\Delta\lambda)\}\phi_3(\Delta\lambda),$$

in which

$$\phi_1(\lambda') \equiv \{\alpha_{p+1}\Gamma_p + \alpha_{p-1}\Gamma_{p+1} e^{\lambda'(\Gamma_{p+1} - \Gamma_p)}\},$$

$$\phi_2(\Delta\lambda) \equiv \alpha_{p+1}\left\{\Gamma_p - \Gamma_{p+1}\frac{1 - e^{\Gamma_p \Delta\lambda}}{1 - e^{\Gamma_{p+1}\Delta\lambda}}\right\},$$

$$\phi_3(\Delta\lambda) \equiv e^{-\Gamma_p \Delta\lambda}\left\{\frac{1 - e^{\Gamma_{p+1}\Delta\lambda}}{\Gamma_{p+1}}\right\}.$$

If the parameter ν lies on a regular path, the relation (13.8) with $k = p$ and $j = 0$, insures the existence of a positive constant ρ which is such that $|\phi_1(\lambda')| > \rho$ for all choices of the root λ' and all positions of ν on the path. The functions $\phi_2(\Delta\lambda)$ and $\phi_3(\Delta\lambda)$ are analytic for small values of $\Delta\lambda$, and vanish at $\Delta\lambda = 0$. A positive constant ϵ may be determined, therefore, such that when $|\Delta\lambda| = \epsilon$, then $|\phi_2(\Delta\lambda)| \leqq \rho/2$, and $|\phi_3(\Delta\lambda)| \neq 0$. It is clear, then, that on the circle

$$\lambda = \lambda' + \Delta\lambda, \qquad |\Delta\lambda| = \epsilon, \tag{15.2}$$

the function $\phi(\lambda)$ fulfills a relation

$$(15.3)\qquad |\phi(\lambda)| > M,$$

with some positive constant M.

Let the function $\psi(\lambda)$ be defined by the formula

$$(15.4)\qquad \begin{aligned}\psi(\lambda) \equiv \{A_p(\lambda, \nu) - \beta_p\nu\}e^{-\lambda\Gamma_p} - \{A_{p+1}(\lambda, \nu) - \alpha_{p+1}\} + A_{p+2}(\lambda, \nu)e^{\lambda\Gamma_{p+1}} \\ - \{A_{p+1}(\lambda, \nu) - \alpha_{p+1}\}e^{\lambda(\Gamma_{p+1}-\Gamma_p)}\}.\end{aligned}$$

On the circle (15.1) an alternative expression for $\psi(\lambda)$ is

$$\begin{aligned}\psi(\lambda) = A_{p+2}(\lambda, \nu)e^{\lambda\Gamma_{p+1}} + \left\{\eta_p \frac{\alpha_{p+1}}{\beta_p} e^{-\Gamma_p\Delta\lambda} - \eta_{p+1} - \beta_{p+1}\nu\right\} \\ + \left\{\eta_p \frac{\alpha_{p+1}}{\beta_p} e^{-\Gamma_{p+1}\Delta\lambda} - \eta_{p+1} - \beta_{p-1}\nu\right\}e^{\lambda(\Gamma_{p+1}-\Gamma_p)},\end{aligned}$$

and from this it may be seen that

$$(15.5)\qquad |\psi(\lambda)| < M,$$

provided N is sufficiently large and $|\nu|$ suitably small, since the exponential exp $\{\lambda(\Gamma_{p+1}-\Gamma_p)\}$ is bounded in the region $S_p(N)$, and the functions exp $\{\lambda\Gamma_{p+1}\}$ and $\eta_i(\lambda, \nu)$, are then arbitrarily small.

By the deductions of §12, both the characteristic values and the roots of the equation (12.16) in the region $S_p(N)$, are represented with an arbitrary degree of accuracy at any specific value of ν ($\neq 0$) by the set of points (12.14) with $k=p$, $j=0$, if the value of N is sufficiently large. Each characteristic value thus corresponds to and is represented in an obvious sense by the respective root λ'. Since with the definitions (15.1) and (15.4) the equation (8.2) is the characteristic equation, it follows from the relations (15.3) and (15.5) by reasoning which is now familiar, that the circle (15.2) contains and retains a characteristic value within it, and therefore that λ' continues to represent its associated characteristic value, so long as it remains in the region $S_p(N)$.

By the formal interchange of the symbols α_{p-1}, β_{p-1}, and Γ_p with α_{p+1}, β_{p+1}, and Γ_{p+1} respectively, the deductions given above may be adapted to the consideration of the characteristic values and roots λ' in the region $S_{p-1}(N)$. Since it was found in §14, that any root λ' which at an initial value of ν lies in the domain comprised of the regions $S_{p-1}(N)$ and $S_p(N)$ remains in this domain as $\nu \to 0$, the asymptotic representability of the characteristic values in the half-plane $S_{p-1,p}$ by means of the roots of the equation (12.16) as $\nu \to 0$, has been established. In particular, therefore, every characteristic value λ_{lm} approaches a finite limit as $\nu \to 0$, and $|\lambda_{lm}|$ is subject to a lower bound, which is indefinitely large with the index m and is independent of the parameter ν.

Since as $\nu \to 0$ along a regular path each circle (15.2) contains precisely one characteristic value, it is a particular consequence that all such values lying outside of some circle centered at $\lambda = 0$ are simple, and that multiplicity is accordingly possible at most in the instance of members of the finite set which lies within such a circle. The following consideration shows, therefore, that multiplicities of the characteristic values may be wholly obviated for values of ν different from zero by an appropriate choice of the path of ν.

Within any circle about the origin, the determinant $D(\lambda, \nu)$ given by the formula (5.15), and $D_\lambda(\lambda, \nu)$, its partial derivative as to λ, are analytic functions of λ and polynomials in ν. Their ν eliminant, therefore, has at most a finite number of zeros within the circle, and these zeros correspond through the characteristic equation (5.7) to the values of ν for which a multiple characteristic value is possible. Inasmuch as these values of ν are thus also finite in number they may,—except for $\nu = 0$, if that is among them,—be avoided by the choice of the path of ν. It will be assumed in the following that any path of ν that is brought into question does avoid these points. The relation

$$(15.6) \qquad D_\lambda(\lambda, \nu) \neq 0, \qquad \text{for } \nu \neq 0,$$

is then fulfilled by every characteristic value.

Chapter 5

Sequences of contours in the λ-plane

16. **An ordering of the characteristic values.** Through their designation in the manner $\lambda_{l,m}$ the characteristic values have been grouped into sub-sets which are distinguished by the respective index values $l = 1, 2, 3, 4$. For the continuing discussion advantages no longer subsist in this, and these values may profitably be regarded hence forth as members of a single simple sequence, in which the ordering is specifically such as will be described in the following.

Let δ be chosen as any positive constant which fulfills the relations

$$(16.1) \qquad \delta < \frac{\pi}{18|\Gamma_j|}, \qquad j = 1, 2,$$

and let δ_1 thereupon designate the smaller one of the values

$$(\delta/3)|\Gamma_j|, \qquad j = 1, 2.$$

For those indices l for which $A_l(\lambda, 0) \equiv 0$ the functions $c_l(\nu)$, given by the relations (12.2), are constant multiples of ν, and hence if ν_0 is an initial parameter value (different from zero) on any regular path that lies in a sector $\sum_{\delta_1}$, the relation

$$(16.2) \qquad |\arg c_l(\nu_0) - \arg c_l(\nu)| < \delta_1,$$

is fulfilled along that path. For those indices l for which $A_l(\lambda, 0)\not\equiv 0$, on the other hand, the functions $c_l(\nu)$ approach non-vanishing limits as $\nu\to 0$. There exists, therefore, a neighborhood of the origin in which the oscillations of the respective functions $\arg c_l(\nu)$ remain less than δ_1, and hence if ν_0 is chosen in such a neighborhood the relation (16.2) is fulfilled for all indices l when ν is on the path segment terminated by ν_0 and the origin. It will be supposed in the following that ν_0 is so chosen, and that ν varies on such a path segment. The relations

$$\left|\frac{1}{\Gamma_h}\arg\frac{c_l(\nu_0)}{c_l(\nu)}\right| < \frac{\delta}{3}, \tag{16.3}$$

then maintain for all indices h and l.

Let the characteristic values be ordered now into a simple sequence

$$\lambda_r(\nu), \qquad r = 1, 2, 3, \cdots, \tag{16.4}$$

with an ordering such that at ν_0 their absolute values stand in a non-decreasing succession, that is,

$$|\lambda_r(\nu_0)| \leqq |\lambda_{r+1}(\nu_0)|, \qquad r = 1, 2, 3, \cdots. \tag{16.5}$$

Through the asymptotic relationship (12.12), which maintains at $\nu=\nu_0$, this ordering evidently serves immediately to order also the corresponding points (12.10), at least insofar as those with sufficiently large indices m are concerned, into the sequence

$$\lambda_r^*(\nu), \qquad r = r_1, r_1+1, r_1+2, \cdots. \tag{16.6}$$

Inasmuch as each member of this latter sequence is drawn from one of the four sub-sets (12.10), it is clear that any consecutive five of them must include at least two from some one of the sub-sets. To every sufficiently large index r, therefore, there corresponds some index pair (l, m) such that

$$|\lambda_{r+4}^*(\nu_0)| - |\lambda_r^*(\nu_0)| \geqq |\lambda_{l,m+1}^*(\nu_0)| - |\lambda_{l,m}^*(\nu_0)|. \tag{16.7}$$

In this relation the right-hand member differs by arbitrarily little from the value $2\pi/|\Gamma_l|$, whenever m is large enough, as may be seen from the formula (12.10). It follows in particular, from the relation (16.1), that the left-hand member of the inequality (16.7) exceeds the value 32δ whenever r exceeds some specifiable value, and hence that for every such index r at least one of the differences

$$|\lambda_{r+1+j}^*(\nu_0)| - |\lambda_{r+j}^*(\nu_0)|, \qquad j = 0, 1, 2, 3,$$

exceeds 8δ. It may be asserted, therefore, that there exists an increasing sequence of integers n of which no one exceeds its predecessor by more than four, and for each of which the relation

$$(16.8)\qquad |\lambda_{n+1}^*(\nu_0)| - |\lambda_n^*(\nu_0)| > 8\delta,$$

is valid.

17. **The sequence of contours C_n.** For each index n for which the relation (16.8) is fulfilled, let the circle in the λ-plane which is centered at the origin and of the radius Λ_n, with

$$(17.1)\qquad \Lambda_n = (1/2)\{ |\lambda_{n+1}^*(\nu_0)| + |\lambda_n^*(\nu_0)| \},$$

be designated as the contour C_n. It will be seen at once that

$$(17.2)\qquad \begin{aligned} |\lambda_r^*(\nu_0)| &< \Lambda_n - 4\delta, &&\text{for } r \leqq n,\\ |\lambda_r^*(\nu_0)| &> \Lambda_n + 4\delta, &&\text{for } r > n, \end{aligned}$$

and hence that at $\nu=\nu_0$ no point of the sequence (16.6) lies within a distance of 4δ from any one of the contours C_n. It is to be shown that a succession of points ν_n may be chosen on the path ν, such that $|\nu_n|$ decreases monotonically to zero, and such that no characteristic value lies within a distance of δ from the contour C_n when ν is between ν_0 and ν_n, namely when ν is on the "path segment" (ν_0, ν_n).

If x, y, x_0, and y_0 are any real values, and $z=x+iy$, $z_0=x_0+iy_0$ the relation

$$|z^2| - |z_0^2| = (x^2 - x_0^2) + (y - y_0)(y + y_0),$$

is an obvious one, which leads easily to the inequality

$$(17.3)\qquad \big| |z| - |z_0| \big| \leqq \frac{|x^2 - x_0^2|}{|z| + |z_0|} + |y - y_0|.$$

Let any member of the set (16.6) be chosen, and let the indices (l, m) be determined so that this member is also given by the formula (12.10). Then with

$$(17.4)\qquad \begin{aligned} x &= \frac{1}{|\Gamma_l|}\log\left|\frac{c_l(\nu)}{c_{l+1}(\nu)}\right|,\\ y &= \frac{1}{|\Gamma_l|}\left\{2m\pi i + \arg\frac{c_l(\nu)}{c_{l+1}(\nu)}\right\}, \end{aligned}$$

and $x_0=x(\nu_0)$, $y_0=y(\nu_0)$, the formula (12.10) yields $|z|=|\lambda_r^*(\nu)|$, $|z_0| = |\lambda_r^*(\nu_0)|$.

At $\nu=\nu_0$, the relations

$$(17.5)\qquad \left\{\frac{1}{|\Gamma_l|}\log\left|\frac{c_l(\nu)}{c_{l+1}(\nu)}\right|\right\}^2 < (\delta/3)\Lambda_n, \qquad l = 1, 2, 3, 4,$$

are all fulfilled for every sufficiently large index n. They evidently continue to be fulfilled as ν varies from ν_0, so long as it remains subject to a condition

$$\log\left|\frac{1}{\nu}\right| \leqq a_2\Lambda_n^{1/2}, \tag{17.6}$$

in which a_2 is an appropriately determined positive constant. With each index n concerned there may, therefore, be associated a point ν_n on the path of ν for which the equality in the relation (17.6) applies. The condition (17.5) is then clearly fulfilled when ν is between ν_0 and ν_n, whereas the sequence $|\nu_n|$ converges monotonically to zero, as was asserted above. With ν on the path segment (ν_0, ν_n), the formulas (17.5) and (16.3) show at once that

$$|x^2 - x_0^2| < (\delta/3)\Lambda_n,$$

and

$$|y - y_0| < (2\delta/3).$$

The relation (17.3) accordingly yields the inequality

$$\big|\,|\lambda_r^*(\nu)| - |\lambda_r^*(\nu_0)|\,\big| < \frac{(1/3)\delta\Lambda_n}{|\lambda_r^*(\nu)| + |\lambda_r^*(\nu_0)|} + 2\delta/3, \tag{17.7}$$

and from this, together with the relations (17.2), it may be concluded that

$$\begin{aligned} |\lambda_r^*(\nu)| &< \Lambda_n - 2\delta, &\quad \text{for } r \leqq n,\\ |\lambda_r^*(\nu)| &> \Lambda_n + 2\delta, &\quad \text{for } r > n. \end{aligned} \tag{17.8}$$

By the deductions of §12 the relations (12.11), with any positive ϵ, are fulfilled for all sufficiently large values of N, and maintain while ν fulfills the respective condition (12.5). Since in these deductions the role of ϵ may be taken by the constant δ above, and since the corresponding role of N is then filled by any of the constants $\Lambda_n - 2\delta$ in which n is sufficiently large, it follows that all characteristic values $\lambda_r(\nu)$ which lie outside of the circle $|\lambda| = \Lambda_n - 2\delta$, fulfill a relation

$$|\lambda_r(\nu) - \lambda_r^*(\nu)| < \delta, \tag{17.9}$$

and do so for all values of ν that satisfy the condition

$$|\nu| \geqq e^{-a_1(\Lambda_n - 2\delta)}.$$

Inasmuch as the relation

$$a_1(\Lambda_n - 2\delta) > a_2\Lambda_n^{1/2},$$

is fulfilled for all sufficiently large indices n, this specification upon ν is implied by the condition (17.6). The relation (17.9) thus applies in particular over the path segment (ν_0, ν_n), and from this, together with the inequalities (17.8), it follows at once that on this path segment the relations

$$(17.10)\qquad \begin{aligned} |\lambda_r(\nu)| &< \Lambda_n - \delta, && \text{for } r \leqq n, \\ |\lambda_r(\nu)| &> \Lambda_n + \delta, && \text{for } r > n, \end{aligned}$$

maintain.

No characteristic value comes within a distance of δ of any contour C_n associated with a sufficiently large index n, therefore, while ν varies along its path from ν_0 to ν_n.

18. **The determinant $D(\lambda, \nu)$ on the contour C_n.** As it is expressed through the formula (5.15), the determinant $D(\lambda, \nu)$ is the sum of four terms with coefficients $A_l(\lambda, \nu)$ that are of the structure (12.2), (12.3). It is to be shown that when, with a sufficiently large index n, λ and ν are respectively on the contour C_n, and the path segment (ν_0, ν_n), the functions $B_l(\lambda, \nu)$ defined by the formulas

$$(18.1)\qquad B_l(\lambda, \nu) \equiv \frac{1}{D(\lambda, \nu)} \left\{(-1)^{l+1}\lambda^{\tau_1+\tau_2}c_l(\nu)e^{\lambda(V_l-V_1)}\right\},$$

are bounded uniformly as to n, namely that there exists some constant M, independent of n and l, such that for all indices l

$$(18.2)\qquad |B_l(\lambda, \nu)| < M, \quad \text{for } \lambda \text{ on } C_n, \text{ and } \nu \text{ on } (\nu_0, \nu_n).$$

When $|\lambda| = \Lambda_n$, with a sufficiently large index n, and ν is on (ν_0, ν_n), the values

$$(18.3)\qquad |c_l(\nu)e^{\lambda V_l}|, \qquad l = 1, 2, 3, 4,$$

all differ from zero. As λ traces the circle C_n, each of these values is in its turn the dominant one upon a respective arc of the circle. If this arc upon which the largest of the values (18.3) is that given by the index h, is denoted by $C_n^{(h)}$, the relations

$$(18.4)\qquad \left|\frac{c_l(\nu)e^{\lambda V_l}}{c_h(\nu)e^{\lambda V_h}}\right| \leqq 1, \qquad \text{for } \lambda, \nu, \text{ on } C_n^{(h)} \text{ and } (\nu_0, \nu_n),$$

are fulfilled for all l, and it is accordingly clear that on this arc the index h also marks the dominant one of the functions (18.1). The relations (18.2) will, therefore, be established if it is shown that there exists a constant M such that for every h

$$(18.5)\qquad |B_h(\lambda, \nu)| < M, \quad \text{for } \nu \text{ on } (\nu_0, \nu_n) \text{ and } \lambda \text{ on } C_n^{(h)}.$$

From the relations (12.1), the arc $C_n^{(h)}$ is seen to lie partly in each of the sectors S_h and S_{h-1}. It consists, therefore, of two contiguous arcs which may be conveniently denoted by $C_n^{(h,j)}$, $j=0, 1$, and which lie in the respective regions S_{h-j}. On each of these arcs the inequality (18.5) may be established in the manner of the following. The formulas (18.1), (5.15) and (12.3) yield for the reciprocal of $B_h(\lambda, \nu)$ the expression

$$
(18.6)\qquad \begin{aligned} \frac{1}{B_h(\lambda,\nu)} = 1 + \chi_h - \frac{A_{h+1}}{c_h} e^{\lambda(\nu_{h+1}-\nu_h)} + \frac{A_{h+2}}{c_h} e^{\lambda(\nu_{h+2}-\nu_h)} \\ - \frac{A_{h+3}}{c_h} e^{\lambda(\nu_{h+3}-\nu_h)}. \end{aligned}
$$

By the use of the relations (11.7) the final two terms in this may be written in the form

$$
A_{h+2}\left\{\frac{c_{h+1}e^{\lambda\nu_{h+1}}}{c_h e^{\lambda\nu_h}}\right\}\left\{\frac{1}{c_{h+1}} e^{\lambda\Gamma_{h+1}}\right\} - A_{h+3}\left\{\frac{1}{c_h} e^{\lambda\Gamma_{h+1}}\right\}.
$$

On the arc $C_n^{(h,0)}$, therefore, their sum is arbitrarily small, in virtue of the relation (18.4), and the fact that with λ in the domain $S_l(\Lambda_n)$ and ν on (ν_0, ν_n) the values (12.6) are arbitrarily small. The remaining terms on the right of the relation (18.6) are expressible in the manner

$$
\left\{1 - \frac{c_{h+1}}{c_h} e^{\lambda\Gamma_h}\right\} + \left\{\chi_h - \chi_{h+1}\frac{c_{h+1}e^{\lambda\nu_{h+1}}}{c_h e^{\lambda\nu_h}}\right\}.
$$

In this the first member is identical with the function

$$
(18.7)\qquad \left\{1 - e^{(\lambda-\lambda_{h,n}^*)\Gamma_h}\right\}
$$

because of the formula (12.10), whereas the remaining member is again arbitrarily small in virtue of the relation (18.4) and the fact that the functions $\chi_l(\lambda, \nu)$ approach zero uniformly as $|\lambda| \to \infty$. Since with ν on the path segment (ν_0, ν_n) and λ on the arc $C_n^{(h,0)}$, the value $(\lambda - \lambda_{h,n}^*)\Gamma_h$ is bounded from the multiples of $2\pi i$, uniformly as to n, as was shown in §17, it follows that the function (18.7), and hence the entire right-hand member of the formula (18.6), is uniformly bounded from zero. Thus with a suitable constant M, the relation (18.5) is established insofar as the values of λ on the arcs $C_n^{(h,0)}$ are concerned.

For the discussion relative to the arcs $C_n^{(h,1)}$ the reasoning above may be essentially adapted by the mere interchange of the roles of the third and fifth terms on the right of the formula (18.6). Thus the sum of the third and fourth terms, when written in the form

$$
-A_{h+1}\left\{\frac{1}{c_h} e^{\lambda\Gamma_h}\right\} + A_{h+2}\left\{\frac{c_{h-1}e^{\lambda\nu_{h-1}}}{c_h e^{\lambda\nu_h}}\right\}\left\{\frac{1}{c_{h-1}} e^{\lambda\Gamma_h}\right\},
$$

is seen to be arbitrarily small, since that is true of the expressions (12.6) with $l = h-1$. The remaining terms on the right of the formula (18.6) are expressible in the form

$$
\left\{1 - e^{-(\lambda-\lambda_{h-1,n}^*)\Gamma_{h-1}}\right\} + \left\{\chi_h - \chi_{h-1}\frac{c_{h-1}e^{\lambda\nu_{h-1}}}{c_h e^{\lambda\nu_h}}\right\},
$$

and this is uniformly bounded from zero, as was the value (18.7). The existence of a constant M for which the inequalities (18.5), and hence (18.2), are valid, may thus be regarded as established.

CHAPTER 6

ADJOINT BOUNDARY PROBLEMS

19. **The definition and the solutions of adjoint boundary problems.** In terms of the square matrices which appear as coefficients in the equations (5.2), the set of relations

$$
\begin{aligned}
\mathfrak{z}'(x, \nu) &= -\,\mathfrak{z}(x, \nu)\{\lambda\mathfrak{R}(x) + \mathfrak{Q}(x)\},\\
\mathfrak{z}(0, \nu) &= \mathfrak{b}(\nu)\mathfrak{B}^{(0)}(\lambda, \nu),\\
\mathfrak{z}(1, \nu) &= -\,\mathfrak{b}(\nu)\mathfrak{B}^{(1)}(\lambda, \nu),
\end{aligned}
\tag{19.1}
$$

may be looked upon as constituting a differential system for a pair of vectors $\mathfrak{b}(\nu)$, and $\mathfrak{z}(x, \nu)$, of the form

$$
\begin{aligned}
\mathfrak{b}(\nu) &\equiv (b_1(\nu), b_2(\nu)),\\
\mathfrak{z}(x, \nu) &\equiv (z_1(x, \nu), z_2(x, \nu)).
\end{aligned}
\tag{19.2}
$$

With λ and ν at any specific values, this system will be characterized as the *adjoint* of the respective differential system (5.2)(9). As in the case of this latter, λ and ν are to be considered as complex scalar parameters. The vector $\mathfrak{b}(\nu)$ will be referred to as the *parametric vector*, and of a pair of vectors (19.2) which together satisfy the equations of the system, the vector $\mathfrak{z}(x, \nu)$ will be called the *solution*(10).

Let $\mathfrak{Y}(x, \lambda)$, as heretofore, be any nonsingular solution of the matrix equation (4.1). The general solution of the differential equation of the system (19.1) is then given by the formula

$$
\mathfrak{z}(x, \nu) \equiv \mathfrak{k}(\nu)\mathfrak{Y}^{-1}(x, \lambda),
\tag{19.3}
$$

in which $\mathfrak{k}(\nu)$ is an arbitrary vector independent of x. Upon substituting this form into the boundary relations of the system, the vector $\mathfrak{k}(\nu)$ is found to be subject to the evaluations

$$
\begin{aligned}
\mathfrak{k}(\nu) &= \mathfrak{b}(\nu)\mathfrak{B}^{(0)}(\lambda, \nu)\mathfrak{Y}(0, \lambda),\\
\mathfrak{k}(\nu) &= -\,\mathfrak{b}(\nu)\mathfrak{B}^{(1)}(\lambda, \nu)\mathfrak{Y}(1, \lambda).
\end{aligned}
\tag{19.4}
$$

(9) The comparative structure of the systems is somewhat better shown if the equations (5.2) are written in the form $\mathfrak{y}^1(x, \nu) = \{\lambda\mathfrak{R}(x) + \mathfrak{Q}(x)\}\mathfrak{y}(x, \nu)$, $\mathfrak{B}^{(0)}(\lambda, \nu)\mathfrak{y}(0, \nu) = \mathfrak{a}(\nu)$, $\mathfrak{B}^{(1)}(\lambda, \nu)\mathfrak{y}(1, \nu) = -\mathfrak{a}(\nu)$.

(10) This formulation of the adjoint differential system differs in some relatively minor details from that given by the author in the paper: *The boundary problem of an ordinary linear differential system in the complex domain*, Trans. Amer. Math. Soc. vol. 46 (1939) p. 165. It is obtainable therefrom, however, by setting $m=2$, $\eta_0=\eta_1=0$, $\eta_2=1$, and $\mathfrak{z}^{(2)}(x) \equiv -\mathfrak{z}^{(1)}(x) \equiv \mathfrak{z}(x, \nu)$.

The consistency of these evaluations, as may be seen by the elimination of $\mathfrak{f}(\nu)$, is contingent upon fulfillment of the condition

$$\mathfrak{b}(\nu)\mathfrak{D}(\lambda, \nu) = \mathfrak{o}, \tag{19.5}$$

in which $\mathfrak{D}(\lambda, \nu)$ is precisely the matrix given by the formula (5.6). A solution of the system may, therefore, exist only in association with a parametric vector which satisfies the condition (19.5). Conversely, it is seen at once, every parametric vector which does satisfy this condition has a solution associated with it through the relations (19.4) and (19.3).

The choice $\mathfrak{b}(\nu) = \mathfrak{o}$ obviously satisfies the equation (19.5). It is, however, uniquely associated with the solution $\mathfrak{z}(x, \nu) \equiv \mathfrak{o}$. This solution, which is thus always available, may properly be regarded as trivial, and to bar it from the further considerations the specification

$$\mathfrak{b}(\nu) \neq \mathfrak{o}, \tag{19.6}$$

will be imposed. Under this condition the possibility of fulfilling the relation (19.5), and hence the existence of a solution $\mathfrak{z}(x, \nu)$, is contingent upon the values of λ and ν, and the differential system (19.1) may accordingly be regarded as defining a family of boundary problems, precisely as such a family is defined by the system (5.2). The two families (19.1) and (5.2) will be defined to be *adjoint*.

Under the restriction (19.6), the equation (19.5) is solvable if and only if λ is a value for which the matrix $\mathfrak{D}(\lambda, \nu)$ is singular. Such values of λ are accordingly to be designated as characteristic values of the boundary problem (19.1). Since, as roots of the equation (5.7), they have already been identified as characteristic values of the boundary problem (5.2), it must be concluded that adjoint boundary problems have the same characteristic values. That every such value is of the same index, namely admits the same number of linearly independent solutions, for each of the two boundary problems follows also. For the numbers of linearly independent vectors $\mathfrak{c}^{(r)}(\nu)$ and $\mathfrak{b}^{(r)}(\nu)$ which satisfy the respective equations

$$\begin{aligned} \mathfrak{D}(\lambda_r, \nu)\mathfrak{c}^{(r)}(\nu) &= \mathfrak{o}, \\ \mathfrak{b}^{(r)}\nu\mathfrak{D}(\lambda_r, \nu) &= \mathfrak{o}, \end{aligned} \tag{19.7}$$

at the characteristic value λ_r, are, of course, either both one or both two, according as the rank of the matrix $\mathfrak{D}(\lambda_r, \nu)$ is one or zero. If this rank is zero, it is clear that λ_r must be of multiplicity at least two as a zero of the determinant $D(\lambda, \nu)$. The multiplicity of a characteristic value is, therefore, never exceeded by its index.

The solutions which are associated with any vectors $\mathfrak{c}^{(r)}(\nu)$ and $\mathfrak{b}^{(r)}(\nu)$ fulfilling the relations (19.7), are given, respectively, by the formula

$$\mathfrak{y}^{(r)}(x, \nu) \equiv \mathfrak{Y}(x, \lambda_r)\mathfrak{c}^{(r)}(\nu), \tag{19.8}$$

and by either one of the equivalent formulas

$$(19.9)\qquad \begin{aligned} \mathfrak{z}^{(r)}(x, \nu) &\equiv \mathfrak{b}^{(r)}(\nu)\mathfrak{B}^{(0)}(\lambda_r, \nu)\mathfrak{Y}(0, \lambda_r)\mathfrak{Y}^{-1}(x, \lambda_r), \\ \mathfrak{z}^{(r)}(x, \nu) &\equiv -\,\mathfrak{b}^{(r)}(\nu)\mathfrak{B}^{(1)}(\lambda_r, \nu)\mathfrak{Y}(1, \lambda_r)\mathfrak{Y}^{-1}(x, \lambda_r). \end{aligned}$$

For the purposes of subsequent deductions certain pairings of these solutions are advantageous, and are definable whenever the characteristic value concerned is of an index equal to its multiplicity.

Let d_{ij} designate the elements of the matrix $\mathfrak{D}$, so that

$$(19.10)\qquad \mathfrak{D}(\lambda, \nu) \equiv (d_{ij}(\lambda, \nu)).$$

If λ_r is a characteristic value of the index and multiplicity one, the relations

$$d_{kl}(\lambda_r, \nu) \neq 0, \qquad D_\lambda(\lambda_r, \nu) \neq 0,$$

maintain if the subscripts (k, l) are suitably chosen. The relations (19.7) are then in particular fulfilled by the vectors

$$(19.11)\qquad \begin{aligned} \mathfrak{c}^{(r)}(\nu) &= \frac{1}{d_{kl}(\lambda_r, \nu)}\begin{pmatrix} d_{k2}(\lambda_r, \nu) \\ -\,d_{k1}(\lambda_r, \nu)\end{pmatrix}, \\ \mathfrak{b}^{(r)}(\nu) &= \frac{1}{D_\lambda(\lambda_r, \nu)}(-\,d_{2l}(\lambda_r, \nu),\ d_{1l}(\lambda_r, \nu)), \end{aligned}$$

and neither of these is the zero vector. With the evaluations (19.11), the solutions (19.8), (19.9) will be said to be a *normal pair*. If λ_s is a characteristic value of index and multiplicity two, it may conventionally be regarded as the pair of coincident values λ_s and λ_{s+1}. Since in this case the value

$$\lim_{\lambda\to\lambda_s} \frac{D(\lambda, \nu)}{(\lambda - \lambda_s)^2}$$

is not zero, whereas it is the determinant of the matrix

$$(19.12)\qquad \mathfrak{D}_\lambda(\lambda_s, \nu) \equiv \left(\lim_{\lambda\to\lambda_s} \frac{d_{ij}(\lambda, \nu)}{\lambda - \lambda_s}\right),$$

this matrix is nonsingular. It may then be verified that the determinations

$$(19.13)\qquad \begin{aligned} \mathfrak{c}^{(s)}(\nu) &= \{\mathfrak{D}_\lambda(\lambda_s, \nu)\}^{-1}\begin{pmatrix} 1 \\ 0 \end{pmatrix}, & \mathfrak{b}^{(s)}(\nu) &= (-\,1, 0), \\ \mathfrak{c}^{(s+1)}(\nu) &= \{\mathfrak{D}_\lambda(\lambda_{s+1}, \nu)\}^{-1}\begin{pmatrix} 0 \\ 1 \end{pmatrix}, & \mathfrak{b}^{(s+1)}(\nu) &= (0, -\,1), \end{aligned}$$

fulfill the relations (19.7), respectively for $r=s$ and $r=s+1$. With each of them the respective solutions (19.8), (19.9) will also be said to be a *normal pair*. No normal pairing of solutions will be defined in the instance of characteristic values whose multiplicities and indices are not equal.

20. **The generalized relation of bi-orthogonality.** Under the normalization of the given boundary problem in §3, and the construction of its imbedding family in §5, each element $v_{ik}(\lambda, \nu)$ of the matrices $\mathfrak{V}^{(0)}(\lambda, \nu)$, $\mathfrak{V}^{(1)}(\lambda, \nu)$ is a polynomial in λ of maximum degree τ_i, and $\tau_1 \geqq \tau_i$. It is readily seen in virtue of this, that the equations

$$(20.1)\qquad \mathfrak{V}^{(h)}(\lambda', \nu) = \mathfrak{V}^{(h)}(\lambda'', \nu) - (\lambda'' - \lambda') \sum_{l=0}^{\tau_1-1} (\lambda'')^l \mathfrak{V}^{(h,l)}(\lambda', \nu), \qquad h = 0, 1,$$

implicitly define the matrices which are therein designated by $\mathfrak{V}^{(h,l)}(\lambda', \nu)$, and that any element $v_{ij}^{(h,l)}(\lambda', \nu)$ of such a matrix is a polynomial in λ' of degree at most $\tau_i - l - 1$ if $\tau_i - l - 1 \geqq 0$, and vanishes identically if $\tau_i - l - 1 < 0$.

The relation

$$\int_0^1 \{\mathfrak{z}(x)\mathfrak{Y}'(x) + \mathfrak{z}'(x)\mathfrak{Y}(x)\} dx = \mathfrak{z}(1)\mathfrak{Y}(1) - \mathfrak{z}(0)\mathfrak{Y}(0),$$

is an evident identity. If the vector $\mathfrak{z}$ and the matrix $\mathfrak{Y}$ involved in it are taken respectively to be any solution $\mathfrak{z}^p(x, \nu)$ and the matrix $\mathfrak{Y}(x, \lambda)$ which appears in the formula (19.8), the equations (19.1) and (4.1) may be used to give the resulting equality the form

$$(20.2)\qquad \begin{aligned} &(\lambda - \lambda_p) \int_0^1 \mathfrak{z}^{(p)}(x, \nu)\mathfrak{R}(x)\mathfrak{Y}(x, \lambda) dx \\ &\qquad + \mathfrak{b}^{(p)}(\nu)\{\mathfrak{V}^{(0)}(\lambda_p, \nu)\mathfrak{Y}(0, \lambda) + \mathfrak{V}^{(1)}(\lambda_p, \nu)\mathfrak{Y}(1, \lambda)\} = 0. \end{aligned}$$

In this expression the matrices $\mathfrak{V}^{(h)}(\lambda_p, \nu)$ may be replaced by their equivalents as given by the formulas (20.1) with $\lambda' = \lambda_p$ and $\lambda'' = \lambda$. The subsequent multiplication on the right by any one of the vectors $(\lambda - \lambda_p)^{-1}\mathfrak{c}^{(q)}(\nu)$, it being assumed that $\lambda \neq \lambda_p$, results then in giving the relation the form

$$(20.3)\qquad \begin{aligned} &\int_0^1 \mathfrak{z}^{(p)}(x, \nu)\mathfrak{R}(x)\mathfrak{Y}(x, \lambda)\mathfrak{c}^{(q)}(\nu) dx \\ &\quad - \mathfrak{b}^{(p)}(\nu) \sum_{l=0}^{\tau_1-1} \lambda^l \{\mathfrak{V}^{(0,l)}(\lambda_p, \nu)\mathfrak{Y}(0, \lambda)\mathfrak{c}^{(q)}(\nu) + \mathfrak{V}^{(1,l)}(\lambda_p, \nu)\mathfrak{Y}(1, \lambda)\mathfrak{c}^{(q)}(\nu)\} \\ &\qquad = \frac{-1}{\lambda - \lambda_p} \mathfrak{b}^{(p)}(\nu)\mathfrak{D}(\lambda, \nu)\mathfrak{c}^{(q)}(\nu). \end{aligned}$$

As $\lambda \to \lambda_q$ it follows from the formula (19.8) and the analyticity of the matrix $\mathfrak{Y}(x, \lambda)$ as to λ, that

$$\lim_{\lambda \to \lambda_q} \mathfrak{Y}(x, \lambda)\mathfrak{c}^{(q)}(\nu) = \mathfrak{y}^{(q)}(x, \nu).$$

It follows similarly from the first of the relations (19.7) that

$$\lim_{\lambda\to\lambda_q} \mathfrak{D}(\lambda, \nu)\mathfrak{c}^{(q)}(\nu) = 0.$$

If $\lambda_p \neq \lambda_q$, therefore, each member of the relation (20.3) approaches a limit as $\lambda \to \lambda_q$, and the limiting form of the relation as it is given below under (20.5)[11] may be regarded as established whenever the solutions $\mathfrak{z}^{(p)}(x, \nu)$ and $\mathfrak{y}^{(q)}(x, \nu)$ are associated with distinct characteristic values. It is to be shown that the relation (20.5) is valid also when $\lambda_p = \lambda_q$, provided the solutions involved are each a member of a normal pair.

When $\lambda_p = \lambda_q$, the limit of the right-hand member of the relation (20.3) is

$$-\mathfrak{b}^{(p)}(\nu)\mathfrak{D}_\lambda(\lambda_q, \nu)\mathfrak{c}^{(q)}(\nu). \tag{20.4}$$

If $\mathfrak{z}^{(p)}(x, \nu)$ and $\mathfrak{y}^{(q)}(x, \nu)$ are not members of the same normal pair, the characteristic value in question is of the index two, and by the convention adopted in §19, $p \neq q$. The vectors $\mathfrak{b}^{(p)}(\nu)$ and $\mathfrak{c}^{(q)}(\nu)$ are in this case evaluated by the formulas (19.13), with (p, q) identified either with $(s, s+1)$ or with $(s+1, s)$. Under either alternative it is found directly that the limit (20.4) is 0, and hence that the relation (20.5) below again maintains.

When $p = q$ the limit (20.4) is easily found to be 1 if the characteristic value is of the index two. The vectors $\mathfrak{b}^{(p)}(\nu)$ and $\mathfrak{c}^{(q)}(\nu)$ are then as given by the formulas (19.13) either with $p = q = s$, or with $p = q = s+1$, and the result is immediate. If the characteristic value is of the index one, the expression (20.4) for the limit is conveniently replaced by

$$-\mathfrak{b}^{(p)}(\nu)\mathfrak{D}_\lambda(\lambda_p, \nu)\mathfrak{c}^{(p)}(\nu) - \frac{d'_{kl}(\lambda_p, \nu)D(\lambda_p, \nu)}{d_{kl}(\lambda_p, \nu)D_\lambda(\lambda_p, \nu)},$$

which is its equivalent, since $D(\lambda_p, \nu) = 0$. The vectors $\mathfrak{b}^{(p)}(\nu)$ and $\mathfrak{c}^{(p)}(\nu)$ are in this instance evaluated by the formulas (19.11), and with these values the limit in question is found, as has been stated, to be 1.

The solutions of adjoint boundary problems which are members of normal pairs thus fulfill the relations

$$\begin{aligned} &\int_0^1 \mathfrak{z}^{(p)}(x, \nu)\mathfrak{N}(x)\mathfrak{y}^{(q)}(x, \nu)dx \\ &\quad - \mathfrak{b}^{(p)}(\nu)\sum_{l=0}^{\tau_1-1} \lambda_p^l\{\mathfrak{B}^{(0,l)}(\lambda_p, \nu)\mathfrak{y}^{(q)}(0, \nu) + \mathfrak{B}^{(1,l)}(\lambda_p, \nu)\mathfrak{y}^{(q)}(1, \nu)\} = \delta_{pq}. \end{aligned} \tag{20.5}$$

It will be observed that in the absence of the indicated sums in their left-hand members, these relations reduce to the expression of a familiar property of weighted bi-orthogonality of the solutions involved. This reduction evidently maintains whenever $\tau_1 = 0$, namely whenever the boundary problem given is

(11) In which $\delta_{pq} = 0$ if $p \neq q$, $\delta_{pp} = 1$.

one in which the boundary relations are independent of the parameter λ. The relations (20.5) may, therefore, be looked upon as generally expressive of a property of the solutions of which bi-orthogonality is a specialization.

CHAPTER 7

EXPANSIONS IN SERIES OF CHARACTERISTIC SOLUTIONS

21. **The formal expansions of arbitrary vectors.** When the parameter ν is on a regular path, and $\nu\neq 0$, all characteristic values, as has been seen, satisfy the relation (15.6), and are therefore simple roots of the characteristic equation. Every such value is, therefore, of the same index as multiplicity, and the characteristic solutions of the adjoint boundary problems accordingly have the property that they may be adjusted to appear without exception as members of normal pairs. It is essential for the continuing discussion that this property be invariably present, namely also at $\nu=0$. Since the boundary problem is then as originally given, the inherence of the property in it must be a matter of assumption, and this it will be made by the following:

HYPOTHESIS 4. *The given boundary problem is one for which every characteristic value is of an index equal to its multiplicity.*

On the basis of this hypothesis it may, and will be, understood in the following, that the designations $\mathfrak{z}^{(p)}(x, \nu)$, $\mathfrak{y}^{(p)}(x, \nu)$, are reserved to solutions of normal pairs.

If with any sequence of scalar coefficients $\alpha_p(\nu)$, $p=1, 2, 3, \cdots$, the series of characteristic solutions in the equation

$$\sum_{p=1}^{\infty} \alpha_p(\nu)\mathfrak{y}^{(p)}(x, \nu) = \mathfrak{f}(x, \nu), \tag{21.1}$$

is convergent uniformly on the interval $0\leqq x \leqq 1$, and defines there the vector $\mathfrak{f}(x, \nu)$ as shown, and if over and above that the related series in the equations

$$\begin{aligned} &\sum_{p=1}^{\infty} \alpha_p(\nu)\lambda_p^l\mathfrak{y}^{(p)}(0, \nu) = \mathfrak{f}^{(0,l)}(\nu), \\ &\sum_{p=1}^{\infty} \alpha_p(\nu)\lambda_p^l\mathfrak{y}^{(p)}(1, \nu) = \mathfrak{f}^{(1,l)}(\nu), \qquad l = 1, 2, 3, \cdots, (\tau_1 - 1), \end{aligned} \tag{21.2}$$

also converge and define the indicated vectors $\mathfrak{f}^{(j,l)}(\nu)$, the coefficients in question necessarily fulfill in turn the relations

$$\begin{aligned} \alpha_p(\nu) = &\int_0^1 \mathfrak{z}^{(p)}(\xi, \nu)\mathfrak{R}(\xi)\mathfrak{f}(\xi, \nu)d\xi \\ &- \mathfrak{b}^{(p)}(\nu) \sum_{l=0}^{\tau_1-1} \left\{\mathfrak{B}^{(0,l)}(\lambda_p, \nu)\mathfrak{f}^{(0,l)}(\nu) + \mathfrak{B}^{(1,l)}(\lambda_p, \nu)\mathfrak{f}^{(1,l)}(\nu)\right\}, \\ &\qquad p = 1, 2, 3, \cdots, \end{aligned} \tag{21.3}$$

in which the abbreviations

$$\mathfrak{f}^{(0,0)}(\nu) = \mathfrak{f}(0, \nu), \qquad \mathfrak{f}^{(1,0)}(\nu) = \mathfrak{f}(1, \nu), \tag{21.4}$$

have been used. This may be established, simply by substituting for the respective vectors $\mathfrak{f}(x, \nu)$ and $\mathfrak{f}^{(j,l)}(\nu)$ in the formulas (21.3), their equivalent series (21.1), (21.2), interchanging the orders of integration and summation, and applying then the relations (20.5).

With the formulas (21.3) thus at hand, the stated conditional basis, upon which their relation with the equations (21.1) and (21.2) has been made evident, may be abandoned. If with an arbitrary vector $\mathfrak{f}(x, \nu)$, whose components are integrable as to x, an arbitrary auxiliary set of vectors $\mathfrak{f}^{(j,l)}(\nu)$ is taken to be associated, the formulas (21.3) relate to these vectors a sequence of scalars $\alpha_p(\nu)$ as indicated. With these scalars as coefficients, the series of characteristic solutions

$$\sum_{p=1}^{\infty} \alpha_p(\nu)\mathfrak{y}^{(p)}(x, \nu), \tag{21.5}$$

is formally determined, and will be referred to in short as an *expansion of the vector* $\mathfrak{f}(x, \nu)$. Inasmuch as this definition of an expansion is wholly formal, the question of the convergence of such an expansion must manifestly be regarded as an open one. More generally, the amenability of any given expansion to evaluation by "means of summability" of any specific type would be a matter calling for investigation, as would also all questions hinging upon the relation which any value thus conventionally assigned to an expansion may bear to the original generating vector $\mathfrak{f}(x, \nu)$.

22. **The expansions as series of residues.** In terms of any analytic nonsingular solution of the matrix equation (4.1), and the corresponding matrix $\mathfrak{D}(\lambda, \nu)$ given by the relation (5.6), let $\mathfrak{G}(x, \xi, \lambda, \nu)$, which is to be known as the "Green's" matrix, be defined by the formulas

$$\begin{aligned} \mathfrak{G}(x, \xi, \lambda, \nu) &\equiv \mathfrak{Y}(x, \lambda)\mathfrak{D}^{-1}(\lambda, \nu)\mathfrak{W}^{(0)}(\lambda, \nu)\mathfrak{Y}(0, \lambda)\mathfrak{Y}^{-1}(\xi, \lambda), \\ &\qquad\qquad \text{for } 0 \leqq \xi \leqq x, \\ \mathfrak{G}(x, \xi, \lambda, \nu) &\equiv -\mathfrak{Y}(x, \lambda)\mathfrak{D}^{-1}(\lambda, \nu)\mathfrak{W}^{(1)}(\lambda, \nu)\mathfrak{Y}(1, \lambda)\mathfrak{Y}^{-1}(\xi, \lambda), \\ &\qquad\qquad \text{for } x < \xi \leqq 1. \end{aligned} \tag{22.1}$$

At any set of arguments (x, ξ) this matrix, as a function of λ, is analytic except at the characteristic values, where singularities are introduced through at least some of the elements of the matrix $\mathfrak{D}^{-1}(\lambda, \nu)$. These singularities are poles, as may be seen from the formula

$$\mathfrak{D}^{-1}(\lambda, \nu) \equiv \begin{pmatrix} d_{22}(\lambda, \nu)/D(\lambda, \nu) & -d_{12}(\lambda, \nu)/D(\lambda, \nu) \\ -d_{21}(\lambda, \nu)/D(\lambda, \nu) & d_{11}(\lambda, \nu)/D(\lambda, \nu) \end{pmatrix}, \tag{22.2}$$

and moreover poles of the first order whenever, as is here the case, each characteristic value is of a multiplicity equal to its index. The residues are therefore non-vanishing, and may, in the case of the Green's matrix, be computed from either one of the formulas (22.1), since the difference of the right-hand members of these formulas is the matrix $\mathfrak{Y}(x, \lambda)\mathfrak{Y}^{-1}(\xi, \lambda)$, and is thus analytic. With the notational choice of the prefix "res_p" to indicate, for any matrix to which it is applied, the residue at the characteristic value λ_p, it follows, therefore, from the formulas (22.1) that

$$(22.3)\quad \text{res}_p\, \mathfrak{G}(x, \xi, \lambda, \nu) = \mathfrak{Y}(x, \lambda_p)\{\text{res}_p\, \mathfrak{D}^{-1}(\lambda, \nu)\}\mathfrak{B}^{(0)}(\lambda_p, \nu)\mathfrak{Y}(0, \lambda_p)\mathfrak{Y}^{-1}(\xi, \lambda_p).$$

With any choice of the characteristic value λ_p, and with λ distinct from it but in a suitably small neighborhood of it, the identity

$$(\lambda - \lambda_p)\mathfrak{D}^{-1}(\lambda, \nu) \equiv \left\{\frac{1}{\lambda - \lambda_p}\mathfrak{D}(\lambda, \nu)\right\}^{-1},$$

is an obvious one. Its limiting form as $\lambda \to \lambda_p$ is contingent upon the multiplicity of this characteristic value. If λ_p is simple, the formula (22.2) shows the limit to be

$$\text{res}_p\, \mathfrak{D}^{-1}(\lambda, \nu) = \frac{1}{D_\lambda(\lambda_p, \nu)}\begin{pmatrix} d_{22}(\lambda_p, \nu) & -d_{12}(\lambda_p, \nu) \\ -d_{21}(\lambda_p, \nu) & d_{11}(\lambda_p, \nu)\end{pmatrix},$$

whereas it may be seen directly when the characteristic value is multiple, say when $\lambda_p = \lambda_{p+1}$, that the limit is

$$\text{res}_p\, \mathfrak{D}^{-1}(\lambda, \nu) = \{\mathfrak{D}_\lambda(\lambda_p, \nu)\}^{-1}.$$

With these two alternatives there are associated respectively the formulas (19.11) and (19.13), and from them it may be verified that

$$(22.4)\qquad \text{res}_p\, \mathfrak{D}^{-1}(\lambda, \nu) = \sum_{\lambda_h = \lambda_p}\{-\mathfrak{c}^{(h)}(\nu)\mathfrak{b}^{(h)}(\nu)\}\text{[12]},$$

in which the sum indicated upon the right consists of the single term for which $h = p$, or of the pair of terms for which $h = p, p+1$, according as λ_p is simple, or $\lambda_p = \lambda_{p+1}$. The substitution of the result (22.4) into the formula (22.3) leads, in virtue of the relations (19.8) and (19.9), to the conclusion that

$$(22.5)\qquad \text{res}_p\, \mathfrak{G}(x, \xi, \lambda, \nu) = \sum_{\lambda_h = \lambda_p}\{-\mathfrak{y}^{(h)}(x, \nu)\mathfrak{z}^{(h)}(\xi, \nu)\}.$$

Consider now, in the case of any expansion (21.5), the term, or pair of terms, associated with any characteristic value. Since $\alpha_p(\nu)$ is a scalar, and

[12] The vectors are to be regarded as matrices for the purposes of the multiplications indicated. Thus $\{\mathfrak{c}^{(h)}(\nu)\mathfrak{b}^{(h)}(\nu)\}$ is a square matrix.

because of the evaluation (19.8), the formula (21.3) leads directly to the equation

$$\alpha_p(\nu)\mathfrak{y}^{(p)}(x, \nu) = \int_0^1 \mathfrak{y}^{(p)}(x, \nu)\mathfrak{z}^{(p)}(\xi, \nu)\mathfrak{R}(\xi)\mathfrak{f}(\xi, \nu)d\xi$$

$$- \mathfrak{Y}(x, \lambda_p)\mathfrak{c}^{(p)}(\nu)\mathfrak{b}^{(p)}(\nu)\sum_{l=0}^{r_1-1}\{\mathfrak{B}^{(0,l)}(\lambda_p, \nu)\mathfrak{f}^{(0,l)}(\nu) + \mathfrak{B}^{(1,l)}(\lambda_p, \nu)\mathfrak{f}^{(1,l)}(\nu)\},$$

and this, together with the results (22.4) and (22.5), yields the relation

$$\begin{aligned}\sum_{\lambda_h=\lambda_p} \alpha_h(\nu)\mathfrak{y}^{(h)}(x, \nu) &= -\int_0^1 \operatorname{res}_p \mathfrak{G}(x, \xi, \lambda, \nu)\mathfrak{R}(\xi)\mathfrak{f}(\xi, \nu)d\xi \\ &+ \operatorname{res}_p \left\{\mathfrak{Y}(x, \lambda)\mathfrak{D}^{-1}(\lambda, \nu)\sum_{l=0}^{r_1-1}[\mathfrak{B}^{(0,l)}(\lambda, \nu)\mathfrak{f}^{(0,l)}(\nu) + \mathfrak{B}^{(1,l)}(\lambda, \nu)\mathfrak{f}^{(1,l)}(\nu)]\right\}.\end{aligned} \tag{22.6}$$

The terms of any expansion (21.5) are thus expressible as residues in the complex plane. It follows from this, of course, that any finite set of such terms may be summed by a suitably designed contour integral as to λ, the contour of integration being chosen to avoid the characteristic values, and to enclose precisely those which are associated with the terms of the set in question. In §17 a certain infinite sequence of contours C_n was defined, any one of the sequence, C_n, enclosing precisely those characteristic values λ_p for which $|\lambda_p| \leqq |\lambda_n|$. If for the values of n there concerned, the initial partial sums of the expansion (21.5) are denoted by $\mathfrak{s}(x, \nu, n)$, in the manner

$$\mathfrak{s}(x, \nu, n) \equiv \sum_{p=1}^{n} \alpha_p(\nu)\mathfrak{y}^{(p)}(x, \nu), \tag{22.7}$$

it follows that these sums are evaluated respectively by the formulas

$$\begin{aligned}\mathfrak{s}(x, \nu, n) &= \frac{-1}{2\pi i}\int_0^1\int_{C_n} \mathfrak{G}(x, \xi, \lambda, \nu)\mathfrak{R}(\xi)\mathfrak{f}(\xi, \nu)d\lambda d\xi \\ &+ \sum_{l=0}^{r_1-1}\frac{1}{2\pi i}\int_{C_n} \mathfrak{Y}(x, \lambda)\mathfrak{D}^{-1}(\lambda, \nu)[\mathfrak{B}^{(0,l)}(\lambda, \nu)\mathfrak{f}^{(0,l)}(\nu) + \mathfrak{B}^{(1,l)}(\lambda, \nu)\mathfrak{f}^{(1,l)}(\nu)]d\lambda.\end{aligned} \tag{22.8}$$

23. **On matters of convergence, divergence, and summability.** When the parameter ν is on a regular path, and $\nu \neq 0$, the formulas (21.3) associate with any suitable vector $\mathfrak{f}(x, \nu)$ an expansion (21.5) in solutions of a boundary problem of the regular type. Such expansions, both in the vector form here in question, and in the alternative scalar form[13], are familiar, and it is known

---

[13] For a discussion of the relations between the scalar and vector formulations, cf. the author's paper, *The expansion problem in the theory of ordinary linear differential systems of the second order*, Trans. Amer. Math. Soc. vol. 31 (1929) p. 887.

that their properties are broadly exemplified by those of the classical Fourier's series. In particular, if x is an interior point of the basic interval, and if in some neighborhood of it the components of the generating vector $\mathfrak{f}(x, \nu)$ are of bounded variation, the expansion converges at this point to the average value, namely

$$\lim_{n\to\infty} \mathfrak{s}(x, \nu, n) = (1/2)\{\mathfrak{f}(x+, \nu) + \mathfrak{f}(x-, \nu)\}, \qquad \nu \neq 0(^{14}). \tag{23.1}$$

These statements, on the other hand, do not ordinarily apply when $\nu=0$. The expansions are then relative to the given boundary problem, which is highly irregular, and little theory of such expansions is known. Indeed, as to boundary problems of the second order—the only ones here immediately pertinent—all highly irregular cases that have been analyzed at all are subsumable in the scalar form

$$\begin{aligned} y''(x) - (2\lambda \cos p\pi/q)y'(x) + \lambda^2 y(x) &= 0, \\ (1-a)y(0) + ay'(0) &= 0, \\ b_1y(0) + b_2y'(0) + b_3y(1) + b_4y'(1) &= 0, \end{aligned} \tag{23.2}$$

with constant coefficients, and in particular with a equal to either 0 or 1, and with p and q relatively prime integers([15]). Moreover, definitive results (uniform convergence), even for the expansions based upon these restricted systems, have been obtained only for highly specialized generating functions, specifically only when these functions are of the structure

$$f(x) \equiv x^{1+a}\phi(x^q),$$

with $\phi(z)$ some analytic function of the complex variable z which is bounded in the circle $|z| < 1$([16]). The disparity between theorems such as these, and those which comprise the theory of expansions relative to regular boundary problems needs no emphasis.

That the expansions associated with highly irregular boundary problems are in general divergent, even when the generating functions are analytic, is observable from the simplest of explicit instances. Thus the expansion generated by the function $f(x)\equiv 1$ relative to the boundary problem (23.2) with $a=0$, $b_4=1$, $b_j=0$, $j=1, 2, 3$, is found to be

$$\sum_{n=1}^{\infty}\left\{e^{x(n\pi-c)\cot c}\,\frac{\sin (n\pi - c)x}{n\pi - c} + e^{-x(n\pi+c)\cot c}\,\frac{\sin (n\pi + c)x}{n\pi + c}\right\}$$

([14]) A proof of this is also implicit in the deductions of Chapter 8 below.

([15]) J. I. Vass, loc. cit.

([16]) For highly irregular boundary problems of order higher than the second, the known expansion theorems are similar and of comparable generality. They refer exclusively to boundary problems in which the differential equation is of the form $y^{(n)}(x)+\{\lambda^n+\sigma(x)\}y(x)=0$, $n\geqq 3$, or some specialization of this form. Cf. J. W. Hopkins and L. E. Ward, loc. cit.

with $c=p\pi/q$. On the interval $0<x<1$ the terms of this series fail to approach zero as $n\rightarrow\infty$. The generating function $f(x)\equiv x^{k+1}$ leads to a similar result whenever k is not an integral multiple of q.

It will be clear, even from these fragmentary citations, that potentialities of fruitfulness in application to expansions of the type (21.5), with generating vectors that are in any broad sense arbitrary, are to be sought only in schemes of evaluation which refer to, and are based upon, some notion more general than that of ordinary convergence. In the following, two such *modes of summation*, each one of a familiar pattern, are to be defined.

To begin with, since the boundary problem originally given is identified with the parameter value $\nu=0$, the expansion generated relative to it by a vector $\mathfrak{f}(x)$, is obtainable from the formula (21.5) by the identification of $\mathfrak{f}(x)$ with $\mathfrak{f}(x, 0)$, and is, thus

$$(23.3)\qquad \sum_{p=1}^{\infty} \alpha_p(0)\mathfrak{y}^{(p)}(x, 0).$$

Now under the hypotheses to which the boundary problems have already been subjected, the characteristic values, and hence also the characteristic solutions, are continuous as functions of ν along any regular path, inclusive of the terminal point $\nu=0$. If the vectors $\mathfrak{f}(x, \nu)$ and $\mathfrak{f}^{(j,l)}(\nu)$ are, therefore, likewise taken to be continuous in ν, and such that

$$(23.4)\qquad \lim_{\nu\to 0} \mathfrak{f}(x, \nu) = \mathfrak{f}(x),$$

it is evident from the formulas (21.3) that each individual term of the expansion (21.5) is continuous and approaches the respective term of the expansion (23.3) as a limit when $\nu\rightarrow 0$. This latter series may, therefore, be regarded as formally given by the expression

$$\lim_{n\to\infty} \lim_{\nu\to 0} \mathfrak{s}(x, \nu, n).$$

The expansion (23.3) is now to be defined as *summable by the "means A" to the value*

$$(23.5)\qquad \lim_{\nu\to 0} \lim_{n\to\infty} \mathfrak{s}(x, \nu, n),$$

if and when with some determination of $\mathfrak{f}(x, \nu)$ *and* $\mathfrak{f}^{(j,l)}(\nu)$ *as vectors continuous in* ν *and fulfilling the relation* (23.4), *the limit* (23.5) *exists.*

The notion of summation basically involved in these means, will be recognized as that which similarly underlies the classical means identified with the names of Abel and Borel. For these latter may be looked upon as evaluating a series

$$(23.6)\qquad \sum_{p=1}^{\infty} u_p(x),$$

respectively, by the limits

$$\lim_{\nu\to 0+} \lim_{n\to\infty} \sum_{p=1}^{n} (1-\nu)^p u_p(x),$$

and

$$\lim_{\nu\to 0+} \lim_{n\to\infty} \sum_{p=1}^{n} \left\{ 1 - e^{1/\nu} \sum_{j=1}^{p} \frac{\nu^{-j}}{j!} \right\} u_p(x).$$

By the use of the relations (23.1) and (23.4) in conjunction with the expression (23.5), it will be evident that the deductions of the preceding sections have effectively established the following facts.

THEOREM. *The expansion* (23.3), *generated by an integrable vector* $\mathfrak{f}(x)$, *is summable by the means A to the value*

$$(23.7) \qquad (1/2)\{\mathfrak{f}(x+) + \mathfrak{f}(x-)\},$$

at any point x of the interval $0<x<1$, *whenever* $\mathfrak{f}(x)$ *is such as to admit at that point of representation in the manner* (23.4), *by a vector* $\mathfrak{f}(x, \nu)$ *which for every* ν *on a regular path other than* $\nu=0$, *fulfills the conditions*

(i) *that it is continuous in* ν;

(ii) *that its expansions relative to regular boundary problems converge to the value*

$$(1/2)\{\mathfrak{f}(x+, \nu) + \mathfrak{f}(x-, \nu)\}.$$

Since the role of $\mathfrak{f}(x, \nu)$ in this theorem may in particular be taken by the vector $\mathfrak{f}(x)$ itself, provided it fulfills the condition (ii), the following specialization of the theorem is evident.

COROLLARY. *The expansion of an integrable vector* $\mathfrak{f}(x)$ *relative to the highly irregular boundary problem is summable by the means A to the value* (23.7), *whenever* $\mathfrak{f}(x)$ *is such that its expansions relative to regular boundary problems converge to that value.*

A second scheme of summation alternative to that described above may be defined in the following manner.

The expansion (23.3) shall be said to be *summable by the "means B" to the value*

$$(23.8) \qquad \lim_{n\to\infty} \mathfrak{s}(x, \nu_n', n),$$

if

(i) *with the role of* $\mathfrak{f}(x, \nu)$ *taken by the vector* $\mathfrak{f}(x)$ *itself, and the vectors* $\mathfrak{f}^{(j,l)}$ *independent of* ν;

(ii) *with the points* $\nu_1', \nu_2', \nu_3', \cdots$ *on some regular path of* ν; and

(iii) *with*

$$\lim_{p\to\infty} \nu_p' = 0,$$

the limit (23.8) *exists.*

The means for summation of an expansion as thus defined may be seen without difficulty to bear in principle a resemblance to the classical means of Cesàro and Riesz. For these latter may be formulated respectively as assigning to a series (23.6) the evaluations

$$\lim_{n\to\infty} \sum_{p=1}^{n} \{1 - (p-1)\nu_n'\} u_p(x), \quad \text{with} \quad \nu_n' = 1/n,$$

and

$$\lim_{n\to\infty} \sum_{p=1}^{n} \{1 - \omega(p)\nu_n'\} u_p(x), \quad \text{with} \quad \nu_n' = 1/\omega(n),$$

the function $\omega(n)$ being positive, increasing and unbounded.

Of the two schemes thus described, the means B may be characterized as providing a subtler mode of summation than the means A, in much the same sense as the means of Cesàro may be regarded as less drastic than those of Borel. It is only consistent with this, that no inference of summability of a highly irregular expansion by the means B is readable from the deductions already made. It is upon this point that the continuing discussion is focused.

Chapter 8

The summability of the expansions by the means B

24. **The formula for the partial sums.** For any index n which identifies a contour of the sequence C_n defined in §17, the terms of an expansion that correspond to the first n characteristic values, are summed by the formula (22.8). The role of this formula in any analysis of the expansion is, of course, a central one; its convergence as $n\to\infty$, with either ν fixed or ν suitably dependent upon n, being tantamount respectively to the convergence or the summability of the expansion. The convergence of these sums, with appropriately disposed parameter values, must, therefore, in due course be considered. Preparatory to this, however, it is to be shown in this section that the formula (22.8) may be expressed in such a manner as to display, among other things, the fact that its elements are bounded as to n when ν is bounded from zero. This is obscured in the formula as it stands, due to the fact that certain of its matrix factors have elements that are polynomials in λ, while others involve exponentials each one of which is clearly unbounded for some range of arg λ as $n\to\infty$.

The Green's matrix has been defined by the formula (22.1). For the purpose of giving alternative expressions for it, let the matrices $\mathfrak{Z}_r$ be defined for all subscripts r by the formulas

$$(24.1)\qquad \begin{aligned} &\mathfrak{J}_1 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \qquad \mathfrak{J}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \\ &\mathfrak{J}_3 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \mathfrak{J}_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \\ &\mathfrak{J}_h = \mathfrak{J}_l, \quad \text{if} \quad h \equiv l \pmod 4), \end{aligned}$$

and let the relations

$$(24.2)\qquad \mathfrak{J}^{(r)} \equiv \begin{cases} \mathfrak{J}_{r+2}, & \text{when} \quad 0 \leqq \xi \leqq x, \\ -\mathfrak{J}_r, & \text{when} \quad x < \xi \leqq 1, \end{cases}$$

specify the matrices $\mathfrak{J}^{(r)}(x, \xi)$. It is then a matter of simple verification that, irrespective of the choice of r, the formulas (22.1) are replaceable by the relation

$$(24.3)\qquad \mathfrak{G} = \mathfrak{Y}(x)\mathfrak{J}^{(r)}\mathfrak{Y}^{-1}(\xi) + \mathfrak{Y}(x)\mathfrak{D}^{-1}\{\mathfrak{B}^{(0)}\mathfrak{Y}(0)\mathfrak{J}_r - \mathfrak{B}^{(1)}\mathfrak{Y}(1)\mathfrak{J}_{r+2}\}\mathfrak{Y}^{-1}(\xi)\text{[17]}.$$

In virtue of the formula (5.6), the relation (24.3) is independent of the choice of $\mathfrak{Y}(x, \lambda)$ as a nonsingular solution of the equation (4.1). This solution may, therefore, be chosen at any value of λ to be one to which the formulas (4.8), (4.9) apply. It will be supposed throughout that the solution $\mathfrak{Y}(x, \lambda)$ is always so chosen. The elements $a_{il}(\lambda, \nu)$, given by the relations (5.3) and (5.9), are then specific, and the matrix $\mathfrak{D}(\lambda, \nu)$ is subject to the formula (5.10).

Let the matrix $\mathfrak{D}_0(\lambda, \nu)$ be specified by the definition

$$(24.4)\qquad \mathfrak{D}_0 \equiv \begin{pmatrix} a_{22} + a_{24}e^{\lambda\Gamma_2} & -a_{12} - a_{14}e^{\lambda\Gamma_2} \\ -a_{21} - a_{23}e^{\lambda\Gamma_1} & a_{11} + a_{13}e^{\lambda\Gamma_1} \end{pmatrix},$$

and let D_0 designate its determinant. A comparison with the formula (5.10) yields then the evaluations

$$(24.5)\qquad \begin{aligned} \mathfrak{D}^{-1} &= \frac{1}{D_0}\mathfrak{D}_0 \cdot (\delta_{ij}\lambda^{-r_i}), \\ D &= \lambda^{r_1+r_2}D_0. \end{aligned}$$

Now it was observed in §20, that the elements of the matrices $\mathfrak{B}^{(h,l)}(\lambda, \nu)$, which occur in the formula (22.8), are polynomials in λ of maximum degree $r_i - l - 1$. In accordance with this, the matrices $\mathfrak{W}^{(h,l)}(\lambda, \nu)$, as defined by the relations

$$(24.6)\qquad \mathfrak{W}^{(h,l)} \equiv \lambda^{l+1}(\delta_{ij}\lambda^{-r_i})\mathfrak{B}^{(h,l)}, \qquad h = 0, 1;\ l = 0, 1, 2, \cdots, (r_1 - 1),$$

have elements that are polynomials in $1/\lambda$, and it is clear that in terms of them

[17] Throughout the remaining discussion the explicit indications of functional arguments will be curtailed in the interest of simplicity in the formulas. Those variables that do not require current attention will therefore frequently be omitted.

$$\mathfrak{D}^{-1}\mathfrak{V}^{(h,l)} = \frac{1}{\lambda^{l+1}D_0}\mathfrak{D}_0\mathfrak{W}^{(h,l)}. \tag{24.7}$$

Since by the relations (4.8), (5.3), (5.9), and (24.5) the evaluations

$$\mathfrak{D}^{-1}\mathfrak{V}^{(0)}\mathfrak{Y}(0) = \frac{1}{D_0}\mathfrak{D}_0(a_{ij}),$$

$$\mathfrak{D}^{-1}\mathfrak{V}^{(1)}\mathfrak{Y}(1) = \frac{1}{D_0}\mathfrak{D}_0(a_{i,j+2})\mathfrak{E}(1),$$

also maintain, it will be recognized that the formula (24.3) is alternatively expressible in the form

$$\mathfrak{G} = \mathfrak{Y}(x)\mathfrak{J}^{(r)}\mathfrak{Y}^{-1}(\xi) + \mathfrak{Y}(x)\frac{1}{D_0}\mathfrak{D}_0\{(a_{ij})\mathfrak{J}_r - (a_{i,j+2})\mathfrak{E}(1)\mathfrak{J}_{r+2}\}\mathfrak{Y}^{-1}(\xi). \tag{24.8}$$

The elements of the matrices $\mathfrak{P}(x, \lambda)$ and $\mathfrak{P}^{-1}(\xi, \lambda)$, which enter into this through the solution $\mathfrak{Y}$ and its inverse, and hence also the elements of the matrices (5.9), are, as has been observed in §5, asymptotically representable by formal power series in negative powers of λ. Through the relations (24.7) and (24.8), the expression of the integrands in the formula (22.8) without the utilization of any positive powers of λ has thus been attained.

For the further analysis of these formulas, let the matrices $\mathfrak{Y}_r(x, \lambda)$ be defined, for all indices r, by the formula

$$\mathfrak{Y}_r(x, \lambda) \equiv \mathfrak{Y}(x, \lambda)\{\mathfrak{J}_r\mathfrak{E}^{-1}(1) + \mathfrak{J}_{r+2}\}. \tag{24.9}$$

The relations

$$\begin{aligned} \mathfrak{Y} &= \mathfrak{Y}_r\{\mathfrak{E}(1)\mathfrak{J}_r + \mathfrak{J}_{r+2}\}, \\ \mathfrak{Y}^{-1} &= \{\mathfrak{J}_{r+2}\mathfrak{E}^{-1}(1) + \mathfrak{J}_r\}\mathfrak{Y}_{r+2}^{-1}, \end{aligned} \tag{24.10}$$

follow at once, and as a consequence the formulas (24.8), (24.7) may be rewritten into the forms

$$\begin{aligned} \mathfrak{G} &= \mathfrak{Y}(x)\mathfrak{J}^{(r)}\mathfrak{Y}^{-1}(\xi) + \mathfrak{Y}_r(x)\mathfrak{U}_r\mathfrak{A}_r\mathfrak{Y}_{r+2}^{-1}(\xi), \\ \mathfrak{Y}(x)\mathfrak{D}\mathfrak{V}^{(h,l)} &= \frac{1}{\lambda^{l+1}}\mathfrak{Y}_r(x)\mathfrak{U}_r\mathfrak{W}^{(h,l)}, \end{aligned} \tag{24.11}$$

in which

$$\begin{aligned} \mathfrak{A}_r &\equiv \{(a_{ij})\mathfrak{J}_r - (a_{i,j+2})\mathfrak{J}_{r+2}\}, \\ \mathfrak{U}_r &\equiv \frac{1}{D_0}\{\mathfrak{E}(1)\mathfrak{J}_r + \mathfrak{J}_{r+2}\}\mathfrak{D}_0. \end{aligned} \tag{24.12}$$

On the contours of the sequence C_n, let the arcs C_{nr} be defined by the relations

$$(24.13)\qquad -\pi/2-\arg\Gamma_{r-1}\leqq\arg\lambda\leqq-\pi/2-\arg\Gamma_r.$$

Any two contiguous arcs of this set comprise a semicircle, and those associated with any four successive values of r constitute a complete contour. Upon associating with each arc C_{nr} the respective evaluations (24.11), the formula (22.8) may now finally be expressed in the form

$$(24.14)\qquad \mathfrak{S}(x,\nu,n)=\mathfrak{S}_0(x,n)+\mathfrak{S}_1(x,\nu,n)+\mathfrak{S}_2(x,\nu,n),$$

with

$$\mathfrak{S}_0(x,n)\equiv\frac{-1}{2\pi i}\sum_{r=1}^{4}\int_0^1\int_{C_{nr}}\{\mathfrak{Y}(x)\mathfrak{J}^{(r)}\mathfrak{Y}^{-1}(\xi)\}\mathfrak{R}(\xi)\mathfrak{f}(\xi)d\lambda d\xi,$$

$$(24.15)\qquad \mathfrak{S}_1(x,\nu,n)\equiv\frac{-1}{2\pi i}\sum_{r=1}^{4}\int_0^1\int_{C_{nr}}\mathfrak{Y}_r(x)\mathfrak{U}_r\mathfrak{A}_r\mathfrak{Y}^{-1}(\xi)\mathfrak{R}(\xi)\mathfrak{f}(\xi)d\lambda d\xi,$$

$$\mathfrak{S}_2(x,\nu,n)\equiv\frac{1}{2\pi i}\sum_{r=1}^{4}\int_{C_{nr}}\mathfrak{Y}_r(x)\mathfrak{U}_r\sum_{l=0}^{r_1-1}\lambda^{-l}\{\mathfrak{W}^{(0,l)}\mathfrak{f}^{(0,l)}+\mathfrak{W}^{(1,l)}\mathfrak{f}^{(1,l)}\}\frac{d\lambda}{\lambda}.$$

Of the matrices which enter into these formulas those designated by $\mathfrak{A}_r$ are shown by the relations (24.12) to involve no exponentials, and to be bounded for all lärge values of $|\lambda|$. The matrices $\mathfrak{U}_r$, on the other hand are less simply constructed. By the formula (18.1) and the second one of the relations (24.5), the equality

$$\frac{1}{D_0}=\frac{(-1)^{r+1}B_r(\lambda,\nu)}{c_r(\nu)}e^{-\lambda(V_r-V_1)},$$

is established, and in virtue of it the matrices in question are found to be explicitly as they are given by the table:

(24.16)

r	1	2	3	4
$\mathfrak{U}_r(\lambda,\nu)$	$\frac{B_1}{c_1(\nu)}\mathfrak{D}_0$	$\frac{-B_2}{c_2(\nu)}\begin{pmatrix}1 & 0\\ 0 & e^{-\lambda\Gamma_1}\end{pmatrix}\mathfrak{D}_0$	$\frac{B_3}{c_3(\nu)}\begin{pmatrix}e^{-\lambda\Gamma_2} & 0\\ 0 & e^{-\lambda\Gamma_1}\end{pmatrix}\mathfrak{D}_0$	$\frac{-B_4}{c_4(\nu)}\begin{pmatrix}e^{-\lambda\Gamma_2} & 0\\ 0 & 1\end{pmatrix}\mathfrak{D}_0$

Consider the exponentials which occur in the elements of any one of the matrices

$$(24.17)\qquad \{\mathfrak{Y}(x)\mathfrak{J}^{(r)}\mathfrak{Y}^{-1}(\xi)\},\qquad \mathfrak{Y}_r(x),\qquad \mathfrak{Y}^{-1}_{r+2}(\xi),\qquad \mathfrak{U}_r.$$

They are in every instance of a form which, in terms of the abbreviations

$$(24.18)\qquad \Gamma_j(x',x'')\equiv\Gamma_j(x')-\Gamma_j(x''),$$

may be written as

(24.19) $$e^{\lambda \Gamma_s(\beta, \alpha)}, \qquad s = r, r + 1,$$

with the arguments α, β each taking some one of the values 0, ξ, x, 1([18]). Moreover, when they are so written, the subscripts s being related to the index r that identifies the arc of integration as is indicated in (24.19), then in every instance the relation $\beta \geqq \alpha$ is fulfilled. It will be shown that because of this every one of the exponentials in question is bounded over the range of integration for which it is involved.

From the formulas (2.5) and (2.7) it follows that when $\beta \geqq \alpha$, then

$$\Gamma_s(\beta, \alpha) = | \Gamma_s(\beta, \alpha) | e^{i \arg \Gamma_s}.$$

Under the substitutions given by the relations

$$\lambda = \mp i\Lambda_n e^{\mp i\theta - i \arg \Gamma_s},$$

the upper or lower signs maintaining according as $s = r$ or $s = r + 1$, the arc C_{nr} corresponds to the range

(24.20) $$0 \leqq \theta \leqq \omega_{r-1},$$

in which the angle ω_{r-1} is that given by the definition (11.8). The evaluations

(24.21) $$| e^{\lambda \Gamma_s(\beta, \alpha)} | = e^{-\Lambda_n | \Gamma_s(\beta, \alpha) | \sin \theta}, \qquad \text{for } \beta \geqq \alpha,$$

then show that on this range the exponentials concerned are bounded as to n. It follows that with the possible exception of $\mathfrak{U}_r$ the matrices (24.17) are all bounded on the respective arcs of integration involved in the formulas (24.15), and that unboundedness can inhere in the matrix $\mathfrak{U}_r$ only through the scalar factors indicated in the table (24.16).

By the relations (18.2) the scalar functions $B_r(\lambda, \nu)$ are shown to be bounded as to n uniformly as to ν, provided the range of ν is restricted to the path segment (ν_0, ν_n) when λ is on the contour C_n. This condition is fulfilled in particular when ν is bounded from zero, and since in this case the boundedness of the coefficients $1/c_r(\nu)$ is also assured, the uniform boundedness of the elements of the matrices $\mathfrak{U}_r(\lambda, \nu)$ follows. If ν is not bounded from zero, on the other hand, this conclusion may not be drawn, since it is the earmark of any highly irregular boundary problem that at least one coefficient $c_r(\nu)$ approaches zero with ν. However, from the formulas (12.2) and (8.4) it is seen that at all events the functions $\nu/c_r(\nu)$ are bounded. It may accordingly be inferred that the elements of the matrices

(24.22) $$\{\nu \mathfrak{U}_r(\lambda, \nu)\},$$

are uniformly bounded for ν on the path segment (ν_0, ν_n) when λ is on the arc C_{nr}.

([18]) It may be noted that $\Gamma_j(x') = \Gamma_j(x', 0)$, and that $\Gamma_j(x', x'') = \Gamma_{j+2}(x'', x')$, since $\Gamma_j(x') = -\Gamma_{j+2}(x')$.

25. **Lemmas.** An analysis of the terms in the relation (24.14), as these are given by the respective formulas (24.15), may be based in large measure upon certain auxiliary deductions, of which some may be regarded as elementary, whereas others are appropriate adaptations of classical convergence theorems. The isolation and specific formulation of these deductions is a matter of evident convenience for their later applications. They will, therefore, be set forth in this section in the form of lemmas, in the interpretation of which it shall be understood that:

(i) Any interval designated by (α, β) is such that $0 \leqq \alpha < \beta \leqq 1$;

(ii) The symbol γ_{nr} is a designation of the semicircle composed of the arcs $C_{n,r-1}$ and C_{nr};

(iii) The range of the index n is the sequence of integers for which the contours C_n have been defined;

(iv) For any value of n the range of the parameter ν is the segment (ν_0, ν_n) of some chosen regular path;

(v) The range of the variable ξ is in every instance an interval (ξ_1, ξ_2) for which $0 \leqq \xi_1 < \xi_2 \leqq 1$;

(vi) Relative to any interval (ξ_1, ξ_2) and any sequence of arcs γ_{nr}, the symbol $\mathfrak{M}(\xi, \lambda, \nu)$ is a generic designation for matrices whose elements $\mathfrak{m}_{ij}(\xi, \lambda, \nu)$ are uniformly bounded, namely which fulfill some set of relations

$$|\mathfrak{m}_{ij}(\xi, \lambda, \nu)| \leqq \rho_{ij}, \tag{25.1}$$

in which the ρ_{ij} are positive constants.

(vii) The symbol $\mathfrak{J}_r(\xi', \xi'', \xi, q)$ is defined by the formula

$$\mathfrak{J}_r(\xi', \xi'', \xi, q) = \int_{\gamma_{nr}} e^{\lambda \Gamma_r(\xi', \xi'')} \mathfrak{M}(\xi, \lambda, \nu) d\lambda/\lambda^q. \tag{25.2}$$

LEMMA 1. *The elements of the matrix*

$$\int_{\gamma_{nr}} \mathfrak{M}(\xi, \lambda, \nu) \frac{d\lambda}{\lambda^{q+1}} \tag{25.3}$$

are bounded uniformly as to ξ, ν and n if $q \geqq 0$, and if $q > 0$ they approach zero uniformly as to ξ and ν, as $n \to \infty$.

The asserted facts are obvious in virtue of the relations (25.1).

LEMMA 2. *If $\beta > \alpha$, the elements of the matrix*

$$\mathfrak{J}_r(\beta, \alpha, \xi, q) \tag{25.4}$$

are bounded uniformly as to ξ, ν and n if $q \geqq 0$, and if $q > 0$ they approach zero uniformly as to ξ and ν, as $n \to \infty$.

Since any arc C_{nr} is identified with the respective range (24.20), the relations

(25.5) $$\sin\theta \geqq \frac{\sin \omega_{r-1-j}}{\omega_{r-1-j}}\theta, \qquad j = 0, 1,$$

maintain on the respective portions $C_{n,r-j}$ of the arc of integration γ_{nr}. By virtue of the evaluations (24.21), it is therefore seen that the elements of the matrix (25.4) are dominated by those of the sum

$$\sum_{j=0}^{1} \Lambda_n^{1-q} \int_0^{\omega_{r-1-j}} \exp\left\{- \Lambda_n \left| \Gamma_r(\beta, \alpha)\right| \frac{\sin \omega_{r-1-j}}{\omega_{r-1-j}}\theta\right\} d\theta(\rho_{ij}).$$

An explicit integration shows these dominant elements to be at all events bounded as to n, and to approach zero if q is positive.

LEMMA 3. *The elements of the matrices*

(25.6) $$\mathfrak{I}_r(\beta, \xi, \xi, q), \quad \textit{with} \quad \xi_2 < \beta,$$

and

(25.7) $$\mathfrak{I}_r(\xi, \alpha, \xi, q), \quad \textit{with} \quad \alpha < \xi_1,$$

are bounded uniformly as to ξ, ν and n if $q \geqq 0$, and if $q>0$ they approach zero uniformly as to ξ and ν, as $n\rightarrow\infty$.

Since in the identity

$$e^{\lambda\Gamma_r(\beta,\xi)} \equiv e^{\lambda\Gamma_r(\beta,\xi_2)}\cdot e^{\lambda\Gamma_r(\xi_2,\xi)},$$

the final exponential is bounded on the arcs γ_{nr}, as is shown by the evaluations (24.21), the matrix (25.6) is in fact of the form (25.4) with $\alpha=\xi_2$. Similarly the matrix (25.7) is of the form (25.4) with $\beta=\xi_1$. The assertions therefore follow from the Lemma 2.

LEMMA 4. *For $\alpha \leqq \xi_1 < \xi_2 \leqq \beta$ the matrices*

(25.8) $$\int_{\xi_1}^{\xi_2} \mathfrak{I}_r(\xi, \alpha, \xi, 1) d\xi,$$

and

(25.9) $$\int_{\xi_1}^{\xi_2} \mathfrak{I}_r(\beta, \xi, \xi, 1) d\xi,$$

approach 0 *as $n\rightarrow\infty$, uniformly as to ξ_1, ξ_2 and ν.*

By virtue of the relations (24.21) and (25.1), the elements of the matrix (25.8) are respectively dominated by those of the sum

$$\sum_{j=0}^{1} \int_{\xi_1}^{\xi_2} \int_0^{\omega_{r-1-j}} \exp\left\{- \Lambda_n \left| \Gamma_r(\xi, \alpha)\right| \sin\theta\right\}(\rho_{ij}) d\theta d\xi.$$

These latter are, however, in turn uniformly dominated by the elements of the matrix sum, which, with any suitably small positive ϵ, is given by the formula

$$\sum_{j=0}^{1}\left\{\int_{\alpha}^{\alpha+\epsilon}\int_{0}^{\omega_{r-1-j}}(\rho_{ij})d\theta d\xi\right.$$
$$\left.+\int_{\alpha+\epsilon}^{\beta}\int_{0}^{\omega_{r-1-j}}\exp\left\{-\Lambda_n\left|\Gamma_r(\alpha+\epsilon,\alpha)\right|\frac{\sin\omega_{r-1-j}}{\omega_{r-1-j}}\theta\right\}(\rho_{ij})d\theta d\xi\right\}.$$

In these matrices the elements of the first two may be made arbitrarily small by the choice of ϵ, and those of the remaining ones are then arbitrarily small when n exceeds some specific value, as is shown by explicit integrations. The convergence of the matrix (25.8) thus follows, and a similar argument establishes the fact for the matrix (25.9).

LEMMA 5. *If* $\mathfrak{K}(\xi, \nu, n)$ *is any matrix such that:*
(i) *its elements are uniformly bounded;*
(ii) *for* $\alpha\leqq\xi_1<\xi_2\leqq\beta$ *the matrix*

$$\int_{\xi_1}^{\xi_2}\mathfrak{K}(\xi,\nu,n)d\xi \tag{25.10}$$

approaches 0 *as* $n\to\infty$, *uniformly as to* ξ_1, ξ_2 *and* ν; *the relation*

$$\lim_{n\to\infty}\int_{\alpha}^{\beta}\mathfrak{K}(\xi,\nu,n)\mathfrak{f}(\xi)d\xi = 0,$$

maintains uniformly as to ν, *for every vector* $\mathfrak{f}(\xi)$ *whose components are integrable over the interval* (α, β).

This is an immediate consequence of a familiar general convergence theorem[19].

LEMMA 6. *If* $\mathfrak{K}(\xi, \nu, n)$ *is any matrix such that:*
(i) *for* $\alpha\leqq\xi_1<\xi_2\leqq\beta$, *the elements of the matrix* (25.10) *are bounded uniformly as to* ξ_1, ξ_2, ν, *and* n;
(ii) *for* $\alpha<\xi_1<\xi_2<\beta$, *the matrix* (25.10) *approaches* 0 *as* $n\to\infty$, *uniformly as to* ν;
(iii) *for* $\alpha<\xi_2<\beta$

$$\lim_{n\to\infty}\int_{\alpha}^{\xi_2}\mathfrak{K}(\xi,\nu,n)d\xi = \mathfrak{L}_\alpha(\nu);$$

(iv) *for* $\alpha<\xi_1<\beta$

---

(19) E. W. Hobson, *The theory of functions of a real variable*, Cambridge University Press, Vol. II, 1926, p. 422.

$$\lim_{n\to\infty}\int_{\xi_1}^{\beta}\mathfrak{K}(\xi, \nu, n)d\xi = \mathfrak{L}_\beta(\nu);$$

the relation

$$\lim_{n\to\infty}\int_{\alpha}^{\beta}\mathfrak{K}(\xi, \nu, n)\mathfrak{f}(\xi)d\xi = \mathfrak{L}_\alpha\mathfrak{f}(\alpha+) + \mathfrak{L}_\beta\mathfrak{f}(\beta-),$$

maintains for every vector $\mathfrak{f}(\xi)$ *whose components are of bounded variation on the interval* (α, β)

This is an evident formulation of a familiar theorem in singular integrals [20].

LEMMA 7. *If* $\mathfrak{K}(\xi, \nu, n)$ *is any matrix which fulfills the specifications* (i) *and* (ii) *of Lemma* 6, *the relation*

$$\lim_{n\to\infty}\int_{\alpha}^{\beta}\mathfrak{K}(\xi, \nu, n)\mathfrak{f}(\xi)d\xi = \mathfrak{o},$$

maintains uniformly as to ν, *for every vector* $\mathfrak{f}(\xi)$ *whose components are of bounded variation on the interval* (α, β), *and for which* $\mathfrak{f}(\alpha+)=\mathfrak{o}$, $\mathfrak{f}(\beta-)=\mathfrak{o}$.

The argument by which the Lemma 6 is established, serves also to prove the assertion here, the conditions (iii) and (iv) of the Lemma 6 being dispensable because of the vanishing limits of the vectors $\mathfrak{f}(\xi)$ concerned, at $\xi=\alpha$ and $\xi=\beta$.

On the basis of these lemmas an analysis of the expressions (24.15) is to be given in the remaining discussion. Consistent with the prime purport of this, which is ultimately to establish summability of the expansions by the means B, it will be assumed henceforth that all vectors $\mathfrak{f}(\xi)$ and $\mathfrak{f}^{(j,l)}$ which are brought into question are independent of ν, and that the vectors $\mathfrak{f}(\xi)$, moreover, all have components that are integrable over the interval (0, 1). The point x at which an expansion is considered will always be regarded as fixed. Since the analysis which applies when x is an end point of the basic interval (0, 1) differs materially from that which is applicable when x is an interior point, these cases will be separately discussed,—the latter in §§26 and 27, and the former finally in §28.

26. **The convergence of the vector $\mathfrak{s}_0(x, n)$, when** $0<x<1$. With the use of the evaluations (24.2) of the matrices $\mathfrak{Z}^{(r)}(x, \xi)$, the formula for the vector $\mathfrak{s}_0(x, n)$, as it is given by the first one of the relations (24.15), is found, after the collection of similar integrals over abutting arcs of integration, to be expressible in the form

$$\mathfrak{s}_0(x, n) = \frac{1}{2\pi i}\int_0^x \mathfrak{K}^{(0)}(\xi, n)\mathfrak{f}(\xi)d\xi + \frac{1}{2\pi i}\int_x^1 \mathfrak{K}^{(1)}(\xi, n)\mathfrak{f}(\xi)d\xi, \tag{26.1}$$

[20] Cf. Hobson, loc. cit. pp. 446–448.

the matrices $\mathfrak{K}^{(h)}(\xi, n)$ being given explicitly by the formulas

$$\mathfrak{K}^{(0)}(\xi, n) \equiv -\sum_{r=1}^{2} \int_{\gamma_{nr}} \mathfrak{Y}(x)\mathfrak{J}_{2r}\mathfrak{Y}^{-1}(\xi)\mathfrak{R}(\xi)d\lambda, \tag{26.2}$$

and

$$\mathfrak{K}^{(1)}(\xi, n) \equiv \sum_{r=3}^{\beta} \int_{\gamma_{nr}} \mathfrak{Y}(x)\mathfrak{J}_{2r}\mathfrak{Y}^{-1}(\xi)\mathfrak{R}(\xi)d\lambda. \tag{26.3}$$

Consider the matrix $\mathfrak{K}^{(0)}(x, \xi)$. Since by the formula (4.8) the matrix $\mathfrak{Y}(x)$ is factorable in the manner $\mathfrak{P}(x)\mathfrak{E}(x)$, whereas the identities

$$\mathfrak{E}(x)\mathfrak{J}_{2r}\mathfrak{E}^{-1}(\xi) \equiv e^{\lambda\Gamma_r(x,\xi)}\mathfrak{J}_{2r}, \qquad r = 1, 2, \tag{26.4}$$

maintain, as may be easily verified, it is seen that the formula (26.2) is given somewhat more explicitly by the form

$$\mathfrak{K}^{(0)}(\xi, n) = -\sum_{r=1}^{2} \int_{\gamma_{nr}} e^{\lambda\Gamma_r(x,\xi)}\mathfrak{P}(x)\mathfrak{J}_{2r}\mathfrak{P}^{-1}(\xi)\mathfrak{R}(\xi)d\lambda. \tag{26.5}$$

Now by the formula (4.9) the matrix $\mathfrak{P}(x)$, and hence also its inverse, differs from the unit matrix by a term which is uniformly of the order of $1/\lambda$. Aside from its scalar exponential factor, the integrand shown in the relation (26.5) is, therefore, of the form

$$\mathfrak{J}_{2r}\mathfrak{R}(\xi) + (1/\lambda)\mathfrak{M}(\xi, \lambda).$$

With the use of the relations

$$\mathfrak{J}_{2r}\mathfrak{R}(\xi) \equiv \Gamma_r'(\xi)\mathfrak{J}_{2r}, \qquad r = 1, 2, \tag{26.6}$$

which follow from the fact that the functions $r_j(\xi)$ which are elements of the matrix $\mathfrak{R}(\xi)$ can also be expressed respectively as $\Gamma_j'(\xi)$, the complete integrands in the formula (26.5) are, therefore, seen to be of the structure

$$e^{\lambda\Gamma_r(x,\xi)}\Gamma_r'(\xi)\mathfrak{J}_{2r} + \mathfrak{J}_r(x, \xi, \xi, 1).$$

An integration with respect to ξ accordingly yields the relation

$$\begin{aligned} \int_{\xi_1}^{\xi_2} \mathfrak{K}^{(0)}(\xi, n)d\xi = \sum_{r=1}^{2}\Bigg\{ \int_{\gamma_{nr}} e^{\lambda\Gamma_r(x,\xi_2)}\mathfrak{J}_{2r}\frac{d\lambda}{\lambda} - \int_{\gamma_{nr}} e^{\lambda\Gamma_r(x,\xi_1)}\mathfrak{J}_{2r}\frac{d\lambda}{\lambda} \\ - \int_{\xi_1}^{\xi_2} \mathfrak{J}_r(x, \xi, \xi, 1)d\xi\Bigg\}. \end{aligned} \tag{26.7}$$

Let x now be fixed upon as any point in the interior of the interval (0, 1) in some neighborhood of which the components of the generating vector $\mathfrak{f}(\xi)$ are of bounded variation. With a suitable determination of ϵ as a positive constant, the neighborhood in question contains the interval $(x-\epsilon, x+\epsilon)$. It will be supposed in the following, that an ϵ has been determined upon which

fulfills this condition. Then, on the one hand, if the points ξ, ξ_1, and ξ_2 lie upon the interval $(0, x-\epsilon)$, the integrals in the formula (26.5) are of the form (25.6) with $q=0$ and $\beta=x$, whereas each integral in the formula (26.7) is either of the form (25.6) with $q=1$, or of the form (25.9). It follows from the Lemmas 3 and 4 that the matrix $\mathfrak{K}^{(0)}(\xi, n)$ fulfills the hypotheses of the Lemma 5, relative to the interval in question, and hence that

$$(26.8) \qquad \lim_{n\to\infty} \int_0^{x-\epsilon} \mathfrak{K}^{(0)}(\xi, n)\mathfrak{f}(\xi)d\xi = \mathfrak{o}.$$

On the other hand, if the points ξ_1 and ξ_2 are taken to lie upon the interval $(x-\epsilon, x)$, it is found similarly by the use of the Lemmas 1, 2, and 4, that the matrix $\mathfrak{K}^{(0)}(\xi, n)$ fulfills the hypotheses of the Lemma 6, with $\alpha=x-\epsilon$, $\beta=x$, and with $\mathfrak{L}_\alpha\equiv\mathfrak{o}$, $\mathfrak{L}_\beta\equiv\pi i\mathfrak{J}$. It follows, therefore, that

$$(26.9) \qquad \lim_{n\to\infty} \int_{x-\epsilon}^{x} \mathfrak{K}^{(0)}(\xi, n)\mathfrak{f}(\xi)d\xi = \pi i\mathfrak{f}(x-).$$

If the consideration is now turned to the matrix $\mathfrak{K}^{(1)}(\xi, n)$ with $x<\xi$, the reasoning given may be essentially repeated. It is found, thus, on the basis of the relations

$$(26.10) \qquad \begin{aligned} \mathfrak{E}(x)\mathfrak{J}_{2r}\,\mathfrak{E}^{-1}(\xi) &\equiv e^{\lambda\Gamma_r(\xi,x)}\mathfrak{J}_{2r}, \\ \mathfrak{J}_{2r}\mathfrak{R}(\xi) &\equiv -\,\Gamma_r'(\xi)\mathfrak{J}_{2r}, \end{aligned} \qquad r = 3, 4,$$

that

$$(26.11) \qquad \lim_{n\to\infty} \int_{x+\epsilon}^{1} \mathfrak{K}^{(1)}(\xi, n)\mathfrak{f}(\xi)d\xi = \mathfrak{o},$$

and that

$$(26.12) \qquad \lim_{n\to\infty} \int_{x}^{x+\epsilon} \mathfrak{K}^{(1)}(\xi, n)\mathfrak{f}(\xi)d\xi = \pi i\mathfrak{f}(x+).$$

The convergence and limiting values of the terms of the formula (26.1) have thus been established, the results admitting of summary in the following form.

The vector $\mathfrak{s}_0(x, n)$ *converges as* $n\to\infty$ *to the value*

$$(26.13) \qquad (1/2)\{\mathfrak{f}(x-) + \mathfrak{f}(x+)\},$$

at every point x which is in the interior of the interval (0, 1) *and in some neighborhood of which the components of the generating vector* $\mathfrak{f}(\xi)$ *are of bounded variation.*

27. **The summability of the expansions at interior points of the interval (0, 1).** By the second one of the formulas (24.15) the product of the vector $\mathfrak{s}_1(x, \nu, n)$ by ν is expressible in the form

$$(27.1) \qquad \nu\mathfrak{s}_1(x, \nu, n) = \int_0^1 \mathfrak{K}(\xi, \nu, n)\mathfrak{f}(\xi)d\xi,$$

with the kernel matrix $\mathfrak{K}(\xi, \nu, n)$ given by the relation

$$(27.2) \qquad \mathfrak{K}(\xi, \nu, n) \equiv \frac{-1}{2\pi i} \sum_{r=1}^{4} \int_{C_{nr}} \mathfrak{Y}_r(x) \{\nu \mathfrak{U}_r(\lambda, \nu)\} \mathfrak{A}_r(\lambda, \nu) \overset{-1}{\mathfrak{Y}}_{r+2}(\xi) \mathfrak{R}(\xi) d\lambda.$$

It was observed in §24, moreover, that each matrix which appears in any integrand of this formula (27.2), has elements that are uniformly bounded over the range of integration concerned, provided the parameter ν is restricted to the respective path segment (ν_0, ν_n). Under this restriction upon ν, which is to be imposed and maintained throughout this discussion, the integrands of the formula (27.2) are thus all of the type $\mathfrak{M}(\xi, \lambda, \nu)$, as that has been defined in §25. A somewhat more explicit determination of the structure of these integrands is requisite, and is obtainable as follows.

By the formula (24.9), the matrices $\mathfrak{Y}_r(x)$ are found to have, for the several indices r, the forms given by the table:

(27.3)

r	$\mathfrak{Y}_r(x, \lambda)$
1	$e^{\lambda\Gamma_1(x,0)}\mathfrak{P}(x)\mathfrak{J}_2 + e^{\lambda\Gamma_2(x,0)}\mathfrak{P}(x)\mathfrak{J}_4$
2	$e^{\lambda\Gamma_3(1,x)}\mathfrak{P}(x)\mathfrak{J}_2 + e^{\lambda\Gamma_2(x,0)}\mathfrak{P}(x)\mathfrak{J}_4$
3	$e^{\lambda\Gamma_3(1,x)}\mathfrak{P}(x)\mathfrak{J}_2 + e^{\lambda\Gamma_4(1,x)}\mathfrak{P}(x)\mathfrak{J}_4$
4	$e^{\lambda\Gamma_1(x,0)}\mathfrak{P}(x)\mathfrak{J}_2 + e^{\lambda\Gamma_4(1,x)}\mathfrak{P}(x)\mathfrak{J}_4$

If these forms are substituted into the relation (27.2), and thereupon the integrals involving any specific exponential over contiguous arcs of integration are collected, it is found that the result may be written in the manner

$$(27.4) \qquad \mathfrak{K}(\xi, \nu, n) = \sum_{r=1}^{2} \int_{\gamma_{nr}} e^{\lambda\Gamma_r(x,0)} \mathfrak{M}(\xi, \lambda, \nu) d\lambda + \sum_{r=3}^{4} \int_{\gamma_{nr}} e^{\lambda\Gamma_r(1,x)} \mathfrak{M}(\xi, \lambda, \nu) d\lambda.$$

Specifically the matrix indicated here by $\mathfrak{M}(\xi, \lambda, \nu)$ over any semicircle γ_{nr}, is identifiable as the product

$$(27.5) \qquad [\mathfrak{P}(x)\mathfrak{J}_{2r}\{\nu\mathfrak{U}_{r-j}\}\mathfrak{A}_{r-j}\overset{-1}{\mathfrak{Y}}_{r+1-j}(\xi)]\mathfrak{R}(\xi) \text{ on the arc } C_{n,r-j}, \qquad j = 0, 1.$$

Consider any matrix $\mathfrak{Z}(x, \lambda, \nu)$ which fulfills a relation

$$(27.6) \qquad \mathfrak{Z}(\xi)\mathfrak{Y}(\xi) \equiv \mathfrak{C},$$

in which $\mathfrak{C}$ is a matrix that is constant as to ξ. It may be deduced from the equation (4.1), then, that $\mathfrak{Z}(\xi)$ is a solution of the adjoint differential equation

$$(27.7) \qquad \mathfrak{Z}'(\xi) = -\mathfrak{Z}(\xi)\{\lambda\mathfrak{R}(\xi) + \mathfrak{Q}(\xi)\},$$

and from this it follows that with any choice of ξ_1 and ξ on the interval (0, 1)

$$\int_{\xi_1}^{\xi} \mathfrak{Z}(\xi)\mathfrak{R}(\xi)d\xi = \frac{-1}{\lambda}\left\{\mathfrak{Z}(\xi) - \mathfrak{Z}(\xi_1) + \int_{\xi_1}^{\xi} \mathfrak{Z}(\xi)\mathfrak{Q}(\xi)d\xi\right\}. \tag{27.8}$$

This formula is applicable to the integration of the relation (27.4). Since the matrix products enclosed within the square brackets in the expressions (27.5) are each of the type prescribed by the designation $\mathfrak{Z}$, the integration of these expressions is given by the formula (27.8). Since, furthermore, each matrix thus designated by $\mathfrak{Z}$ is of the type denotable by $\mathfrak{M}(\xi, \lambda, \nu)$, the entire right-hand member of any resulting relation (27.8) is clearly of the form $(1/\lambda)\mathfrak{M}(\xi, \lambda, \nu)$, uniformly as to the choice of ξ_1, on the interval (0, 1). The formula (27.2) leads, therefore, also to the relations

$$\begin{aligned} \int_{\xi_1}^{\xi_2} \mathfrak{K}(\xi, \nu, n)d\xi = \sum_{r=1}^{2}\int_{\gamma_{nr}} e^{\lambda\Gamma_r(x,0)}\mathfrak{M}(\xi_2, \lambda, \nu)\frac{d\lambda}{\lambda} \\ + \sum_{r=3}^{4}\int_{\gamma_{nr}} e^{\lambda\Gamma_r(1,x)}\mathfrak{M}(\xi_2, \lambda, \nu)\frac{d\lambda}{\lambda}. \end{aligned} \tag{27.9}$$

The integrals in the right-hand members of the formulas (27.4) and (27.9) may now all be recognized as being of the forms (25.4) with q, respectively, equal to either 0 or 1. It follows from this by the Lemma 2 that the matrix (27.2) fulfills the hypotheses of the Lemma 5. By that lemma, then, the relation (27.1) leads to the conclusion that

$$\lim_{n\to\infty} \{\nu\mathfrak{s}_1(x, \nu, n)\} = 0, \tag{27.10}$$

uniformly as to ν on the segment (ν_0, ν_n).

The analysis which has thus been given for the second one of the formulas (24.15) may be applied equally well, and in a wholly similar manner to the third one of these formulas. It yields in this instance an evaluation of the form

$$\begin{aligned} \nu\mathfrak{s}_2(x, \nu, n) = \sum_{l=0}^{r_1-1}\sum_{h=0}^{1}\Bigg\{\sum_{r=1}^{2}\int_{\gamma_{nr}} e^{\lambda\Gamma_r(x,0)}\mathfrak{M}(\lambda, \nu)\frac{d\lambda}{\lambda^{l+1}}\mathfrak{f}^{(h,l)} \\ + \sum_{r=3}^{4}\int_{\gamma_{nr}} e^{\lambda\Gamma_r(1,x)}\mathfrak{M}(\lambda, \nu)\frac{d\lambda}{\lambda^{l+1}}\mathfrak{f}^{(h,l)}\Bigg\}, \end{aligned} \tag{27.11}$$

and inasmuch as each integrand on the right of this equality is of the type (25.4) with $q>0$, it follows from the Lemma 2 that irrespective of the vectors $\mathfrak{f}^{(h,l)}$ involved,

$$\lim_{n\to\infty} \{\nu\mathfrak{s}_2(x, \nu, n)\} = 0, \tag{27.12}$$

uniformly as to ν on the segment (ν_0, ν_n).

In virtue of the conclusions (27.10) and (27.12), it may evidently be in-

ferred that there exists a positive monotonic sequence of constants

$$\epsilon_1, \epsilon_2, \epsilon_3, \cdots$$

which converges to zero, and which is furthermore such that for each index n, and for all values of ν on the respective path segment (ν_0, ν_n), the relations

$$(27.13) \qquad \left| \nu \mathfrak{S}_1(x, \nu, n) + \nu \mathfrak{S}_2(x, \nu, n) \right| < \epsilon_n,$$

are fulfilled. On the path of ν let the sequence of points

$$(27.14) \qquad \nu_1', \nu_2', \nu_3', \cdots$$

be determined now so that

(i) ν_n' lies on the segment (ν_0, ν_n),

(ii)

$$(27.15) \qquad \lim_{n\to\infty} \nu_n' = 0,$$

and

(iii)

$$(27.16) \qquad \lim_{n\to\infty} \frac{\epsilon_n}{\nu_n'} = 0.$$

The determination of such a sequence is clearly possible. Since the inequalities (27.13) are, then, in particular fulfilled when $\nu = \nu_n'$, it follows from them, together with the relation (27.16), that

$$(27.17) \qquad \lim_{n\to\infty} \left\{ \mathfrak{S}_1(x, \nu_n', n) + \mathfrak{S}_2(x, \nu_n', n) \right\} = 0.$$

This result, taken in conjunction with that of §26, permits of the conclusion which follows.

THEOREM. *The expansion generated by an arbitrary integrable vector* $\mathfrak{f}(\xi)$, *and an arbitrary set of associated vectors* $\mathfrak{f}^{(h,l)}$, $h=0, 1; l=1, 2, 3, \cdots, (\tau_1-1)$; *is summable by the means B to the value* $(1/2)\{\mathfrak{f}(x-)+\mathfrak{f}(x+)\}$ *at every interior point of the basic interval in some neighborhood of which the components of the vector* $\mathfrak{f}(\xi)$ *are of bounded variation.*

28. **The summability of the expansions at the end points of the interval.** The reasoning of the two immediately preceding sections depends explicitly upon the relation $0<x<1$, and is essentially inadequate when the point at which the expansion is considered is an end point of the basic interval. At these points, $x=0$, and $x=1$, therefore, distinct considerations are requisite. Such are to be given in the following, it being assumed throughout their course, firstly, that the given boundary problem has, if necessary, been adjusted in the manner described in §3, so that the relations (3.2) maintain; and secondly, that only such expansions are brought into question as have generating vectors $\mathfrak{f}(\xi)$ whose components are of bounded variation in some

right-hand neighborhood of $\xi=0$ and in some left-hand neighborhood of $\xi=1$. A positive constant ϵ may be determined in each case, then, so that the boundedness of variation assumed is maintained over each one of the intervals $(0, \epsilon)$ and $(1-\epsilon, 1)$. It will be supposed that ϵ has been so determined. The pair of intervals $(0, \epsilon)$, $(1-\epsilon, 1)$ will briefly be designated by the symbol Δ, and it shall be understood that the designations $\mathfrak{f}(0)$ and $\mathfrak{f}(1)$ signify the limiting values $\mathfrak{f}(0+)$ and $\mathfrak{f}(1-)$, respectively.

As a function of ξ the Green's matrix $\mathfrak{G}(x, \xi, \lambda, \nu)$ is a solution of the differential equation (27.7), as may be seen at once from its definition (22.1). It is found in virtue of this that the relation

$$-\int_0^1 \mathfrak{G}\mathfrak{R}\mathfrak{f}d\xi = \int_\Delta \{\mathfrak{G}_\xi + \mathfrak{G}\mathfrak{Q}\}\mathfrak{f}\frac{d\xi}{\lambda} - \int_\epsilon^{1-\epsilon} \mathfrak{G}\mathfrak{R}\mathfrak{f}d\xi, \tag{28.1}$$

maintains, and furthermore that every differentiable vector $\mathfrak{g}(\xi)$ satisfies the identity

$$\begin{aligned} 0 = \int_0^1 \frac{\partial}{\partial\xi}\{\mathfrak{G}\mathfrak{g}\}\frac{d\xi}{\lambda} - \int_\Delta \{\mathfrak{G}_\xi\mathfrak{g} + \mathfrak{G}\mathfrak{g}'\}\frac{d\xi}{\lambda} \\ + \int_\epsilon^{1-\epsilon} \{\lambda\mathfrak{G}\mathfrak{R}\mathfrak{g} + \mathfrak{G}\mathfrak{Q}\mathfrak{g} - \mathfrak{G}\mathfrak{g}'\}\frac{d\xi}{\lambda}. \end{aligned} \tag{28.2}$$

The specific vector for which this identity is to be utilized, and to which the designation $\mathfrak{g}(\xi)$ will hereafter be restricted, is the following one:

$$\mathfrak{g}(\xi) \equiv \xi\mathfrak{f}(1) + (1-\xi)\mathfrak{f}(0). \tag{28.3}$$

If the respective members of the identities (28.1) and (28.2) are added, and their sum is integrated over the contour C_n, the result is an evaluation of first integral in the formula (22.8). With a suitable grouping of the terms over the several ranges of integration, that formula is, therefore, found to be expressible in the form

$$\mathfrak{s}(x, \nu, n) = \mathfrak{f}(x) + \sum_{k=0}^{5} \mathfrak{s}_{0k}(x, \nu, n), \tag{28.4}$$

with

$$\begin{aligned} \mathfrak{s}_{00} \equiv -\mathfrak{f}(x) + \frac{1}{2\pi i}\int_0^1 \int_{C_n} \frac{\partial}{\partial\xi}\{\mathfrak{G}\mathfrak{g}\}\frac{d\lambda}{\lambda}d\xi \\ + \frac{1}{2\pi i}\int_{C_n} \mathfrak{Y}(x)\mathfrak{D}^{-1}\{\mathfrak{B}^{(0,0)}(\lambda)\mathfrak{f}(0) + \mathfrak{B}^{(1,0)}(\lambda)\mathfrak{f}(1)\}d\lambda, \end{aligned} \tag{28.5}$$

and

$$\mathfrak{s}_{01} \equiv \sum_{l=1}^{r_1-1} \frac{1}{2\pi i}\int_{C_n} \mathfrak{Y}(x)\mathfrak{D}^{-1}\{\mathfrak{B}^{(0,l)}(\lambda)\mathfrak{f}^{(0,l)} + \mathfrak{B}^{(1,l)}(\lambda)\mathfrak{f}^{(1,l)}\}d\lambda, \tag{28.6}$$

and with the remaining terms expressible through the matrices

$$
(28.7)\qquad
\begin{aligned}
\mathfrak{K}^{(1)}(x, \xi, \nu, n) &\equiv \frac{1}{2\pi i}\int_{C_n} \mathfrak{G}(x, \xi, \lambda, \nu)\frac{d\lambda}{\lambda},\\
\mathfrak{K}^{(2)}(x, \xi, \nu, n) &\equiv \frac{1}{2\pi i}\int_{C_n} \mathfrak{G}_\xi(x, \xi, \lambda, \nu)\frac{d\lambda}{\lambda},\\
\mathfrak{K}^{(3)}(x, \xi, \nu, n) &\equiv \frac{1}{2\pi i}\int_{C_n} \mathfrak{G}(x, \xi, \lambda, \nu)\mathfrak{R}(\xi)d\lambda,
\end{aligned}
$$

by the formulas

$$
(28.8)\qquad
\begin{aligned}
\mathfrak{s}_{02} &= \int_\epsilon^{1-\epsilon} \mathfrak{K}^{(1)}\{\mathfrak{Q}\mathfrak{g} - \mathfrak{g}'\}d\xi,\\
\mathfrak{s}_{03} &= \int_\Delta \mathfrak{K}^{(1)}\{\mathfrak{Q}\mathfrak{f} - \mathfrak{g}'\}d\xi,
\end{aligned}
$$

and

$$
(28.9)\qquad
\begin{aligned}
\mathfrak{s}_{04} &= \int_\Delta \mathfrak{K}^{(2)}\{\mathfrak{f} - \mathfrak{g}\}d\xi,\\
\mathfrak{s}_{05} &= -\int_\epsilon^{1-\epsilon} \mathfrak{K}^{(3)}\{\mathfrak{f} - \mathfrak{g}\}d\xi.
\end{aligned}
$$

Consider the formula (28.5). If the indicated integration as to ξ therein is performed, and the expression $\mathfrak{f}(x)$ is replaced by its equivalent

$$
\frac{1}{2\pi i}\int_{C_n} \mathfrak{f}(x)\frac{d\lambda}{\lambda},
$$

the formula is seen to be alternatively

$$
(28.10)\qquad
\begin{aligned}
\mathfrak{s}_{00} = \frac{1}{2\pi i}\int_{C_n} \{&\mathfrak{G}(x, 1-)\mathfrak{f}(1) - \mathfrak{G}(x, 0+)\mathfrak{f}(0) - \mathfrak{f}(x)\\
&+ \mathfrak{Y}(x)\mathfrak{D}^{-1}[\lambda\mathfrak{B}^{(0,0)}(\lambda)\mathfrak{f}(0) + \lambda\mathfrak{B}^{(1,0)}(\lambda)\mathfrak{f}(1)]\}\frac{d\lambda}{\lambda}\cdot
\end{aligned}
$$

Now the definition (22.1) of the Green's matrix leads to the evaluations

$$
(28.11)\qquad
\begin{aligned}
\mathfrak{G}(0, 0+) &= -\mathfrak{J} + \mathfrak{Y}(0)\mathfrak{D}^{-1}\mathfrak{B}^{(0)}(\lambda),\\
\mathfrak{G}(0, 1-) &= \quad\; - \mathfrak{Y}(0)\mathfrak{D}^{-1}\mathfrak{B}^{(1)}(\lambda),\\
\mathfrak{G}(1, 0+) &= \quad\;\; \mathfrak{Y}(1)\mathfrak{D}^{-1}\mathfrak{B}^{(0)}(\lambda),\\
\mathfrak{G}(1, 1-) &= \mathfrak{J} - \mathfrak{Y}(1)\mathfrak{D}^{-1}\mathfrak{B}^{(1)}(\lambda),
\end{aligned}
$$

whereas it follows from the relations (20.1), with λ' and λ'' identified, respectively, as λ and 0, that

$$
(28.12)\qquad \lambda\mathfrak{B}^{(h,0)}(\lambda) = \mathfrak{B}^{(h)}(\lambda) - \mathfrak{B}^{(h)}(0), \qquad h = 0, 1.
$$

In virtue of these substitutions the formula (28.10) reduces, irrespective of whether $x=0$ or $x=1$, to the form

$$\mathfrak{s}_{00} = \frac{-1}{2\pi i}\int_{C_n} \mathfrak{Y}(x)\mathfrak{D}^{-1}\frac{d\lambda}{\lambda}\{\mathfrak{W}^{(0)}(0)\mathfrak{f}(0) + \mathfrak{W}^{(1)}(0)\mathfrak{f}(1)\},$$

and this leads finally, through the evaluations

$$\mathfrak{Y}(x)\mathfrak{D}^{-1}(\lambda) = \mathfrak{Y}_r(x)\mathfrak{U}_r\{\lambda^{-r_1}\mathfrak{J}_2 + \lambda^{-r_2}\mathfrak{J}_4\}, \qquad r = 1, 2, 3, 4,$$

to the relation

$$(28.13)\quad \nu\mathfrak{s}_{00} = \frac{-1}{2\pi i}\sum_{r=1}^{4}\sum_{j=1}^{2}\int_{C_{nr}} \mathfrak{Y}_r(x)\{\nu\mathfrak{U}_r\}\mathfrak{J}_{2j}\frac{d\lambda}{\lambda^{r_j+1}}\{\mathfrak{W}^{(0)}(0)\mathfrak{f}(0) + \mathfrak{W}^{(1)}(0)\mathfrak{f}(1)\}.$$

The transformation of the formula (28.6) is more direct. Through the mere use of the second one of the equations (24.11), it yields, namely the relation

$$(28.14)\qquad \nu\mathfrak{s}_{01} = \frac{1}{2\pi i}\sum_{l=1}^{r_1-1}\sum_{r=1}^{4}\sum_{h=0}^{1}\int_{C_{nr}} \mathfrak{Y}_r(x)\{\nu\mathfrak{U}_r\}\mathfrak{W}^{(h,l)}\frac{d\lambda}{\lambda^{l+1}}\mathfrak{f}^{(h,l)}.$$

Under the restriction of the parameter ν to the path segment (ν_0, ν_n), each matrix factor which appears in an integrand in the formulas (28.13) and (28.14) has been seen to be uniformly bounded over the range of integration for which it is involved. Due to the relations (3.2), and by the Lemma 1, therefore, it may be concluded that

$$(28.15)\qquad \begin{aligned} &\lim_{n\to\infty}\{\nu\mathfrak{s}_{00}(x, \nu, n)\} = 0,\\ &\lim_{n\to\infty}\{\nu\mathfrak{s}_{01}(x, \nu, n)\} = 0, \end{aligned}$$

uniformly as to ν on the segment (ν_0, ν_n).

On the basis of the first one of the formulas (24.11), and by considerations which are now familiar, it is found that both when $x=0$ and when $x=1$ the definitions (28.7) assure for the matrices $\nu\mathfrak{K}^{(1)}(x, \xi, \nu, n)$ and $\nu\mathfrak{K}^{(2)}(x, \xi, \nu, n)$ the forms

$$(28.16)\quad \nu\mathfrak{K}^{(2-2j)} = \sum_{r=1}^{2}\int_{\gamma_{nr}} e^{\lambda\Gamma_r(1,\xi)}\mathfrak{M}(\xi, \lambda, \nu)\frac{d\lambda}{\lambda^j} + \sum_{r=3}^{4}\int_{\gamma_{nr}} e^{\lambda\Gamma_r(\xi,0)}\mathfrak{M}(\xi, \lambda, \nu)\frac{d\lambda}{\lambda^j},$$

$$j = 0, 1,$$

provided ν is restricted to range between ν_0 and ν_n. For any choice of ξ_1 and ξ_2, moreover,

$$(28.17)\qquad \int_{\xi_1}^{\xi_2} \nu\mathfrak{K}^{(2)}(x, \xi)d\xi = \nu\mathfrak{K}^{(1)}(x, \xi_2) - \nu\mathfrak{K}^{(1)}(x, \xi_1),$$

and

$$\begin{aligned}\int_{\xi_1}^{\xi_2} \nu\mathfrak{K}^{(3)}(x, \xi_s)d\xi = &-\nu\mathfrak{K}^{(1)}(x, \xi_2) + \nu\mathfrak{K}^{(1)}(x, \xi_1) \\ &- \int_{\xi_1}^{\xi_2} \nu\mathfrak{K}^{(1)}(x, \xi)\mathfrak{Q}(\xi)d\xi,\end{aligned} \tag{28.18}$$

the latter one of these relations following from the fact that the matrix $\mathfrak{G}$ satisfies the differential equation (27.7).

The Lemmas 1 and 4 applied to the formula (28.16) with $j=1$, show readily that the matrix $\nu\mathfrak{K}^{(1)}(x, \xi)$ fulfills the hypotheses of the Lemma 5 with $\mathfrak{L}_\alpha(\nu) \equiv \mathfrak{O}$, and $\mathfrak{L}_\beta(\nu) \equiv \mathfrak{O}$, and with (α, β) as any subinterval of the interval $(0, 1)$. Since the vectors $\{\mathfrak{Q}(\xi)\mathfrak{g}(\xi) - \mathfrak{g}'(\xi)\}$ and $\{\mathfrak{Q}(\xi)\mathfrak{f}(\xi) - \mathfrak{g}'(\xi)\}$ have components that are of bounded variation respectively on the interval $(\epsilon, 1-\epsilon)$ and the pair of intervals Δ, it follows from the formulas (28.8) and the Lemma 5 that

$$\begin{aligned}&\lim_{n\to\infty} \{\nu\mathfrak{S}_{02}(x, \nu, n)\} = \mathfrak{o}, \\ &\lim_{n\to\infty} \{\nu\mathfrak{S}_{03}(x, \nu, n)\} = \mathfrak{o},\end{aligned} \tag{28.19}$$

uniformly as to ν on the segment (ν_0, ν_n).

The formulas (28.17) and (28.16) with $j=1$, show readily that the matrix $\nu\mathfrak{K}^{(2)}(x, \xi)$ fulfills the hypotheses of the Lemma 7 relative to the intervals Δ. The formulas (28.18) and (28.16) with $j=0$, show similarly that the matrix $\nu\mathfrak{K}^{(3)}(x, \xi)$ fulfills the hypotheses of the Lemma 5 relative to the interval $(\epsilon, 1-\epsilon)$. In virtue of the formulas (28.9), and the fact that the vector $\{\mathfrak{f}(\xi) - \mathfrak{g}(\xi)\}$ vanishes at $\xi = 0$ and $\xi = 1$, it follows, therefore, lastly that

$$\begin{aligned}&\lim_{n\to\infty} \{\nu\mathfrak{S}_{04}(x, \nu, n)\} = \mathfrak{o}, \\ &\lim_{n\to\infty} \{\nu\mathfrak{S}_{05}(x, \nu, n)\} = \mathfrak{o},\end{aligned} \tag{28.20}$$

uniformly as to ν on the segment (ν_0, ν_n).

The results (28.15), (28.19) and (28.20) evidently insure the existence of a sequence of positive constants $\epsilon_1, \epsilon_2, \epsilon_3, \cdots$ which converges to zero, and which is such that the relations

$$\left| \nu \sum_{k=0}^{5} \mathfrak{S}_{0k}(x, \nu, n) \right| < \epsilon_n, \qquad n = 1, 2, 3, \cdots,$$

maintain, irrespective of how in the nth one of them the value of ν is chosen on the path segment (ν_0, ν_n). In particular, then, these points may be chosen as the respective members of a sequence (27.14) which fulfills the relations (27.15) and (27.16). For such a choice it is clear that

$$\lim_{n\to\infty} \sum_{k=0}^{5} \mathfrak{s}_{0k}(x, \nu_n', n) = \mathfrak{o},$$

and this result yields through the relation (28.4) the following and final conclusion.

THEOREM. *The expansion generated by an arbitrary integrable vector* $\mathfrak{f}(\xi)$ *whose components are of bounded variation in some right-hand neighborhood of the point* $\xi=0$, *and in some left-hand neighborhood of the point* $\xi=1$, *is summable by the means B to the vector* $\mathfrak{f}(x)$ *at the points* $x=0$ *and* $x=1$.

UNIVERSITY OF WISCONSIN,
MADISON, WIS.

ON APPROXIMATING CERTAIN INTEGRALS BY SUMS

BY
C. RAYMOND ADAMS AND ANTHONY P. MORSE

1. Introduction; the main problem and its setting. In a recent paper[1] we showed that if E is a measurable linear set (bounded or unbounded), if $f \in L(E)$, if k is a fixed real number ≥ 1, and if A_n and B_n $(n=1, 2, 3, \cdots)$ are sequences (finite or infinite) of measurable subsets of E satisfying the conditions

$$A_n \subset B_n, \quad 0 < |B_n| \leq k|A_n|, \quad \text{diameter } (B_n) < \delta \qquad (n = 1, 2, 3, \cdots);$$
$$|B_m B_n| = 0 \quad \text{for} \quad m \neq n; \qquad \sum_n B_n = E,$$

then the following relation holds:

$$\int_E f = \lim_{\delta \to 0} \sum_n \left[\int_{A_n} f / |A_n| \right] |B_n|.$$

The present paper is concerned chiefly with the generalization of this result obtained by replacing $f(x)$ by $\phi[f(x)]$. From the statement above it is evident that the *order* of the elements of the sequences A_n and B_n $(n=1, 2, 3, \cdots)$ is *immaterial.* It is therefore desirable to introduce a notation which will be free from any implication of order. To achieve this purpose and to enable us to state our present problem with precision we formulate several definitions as follows.

(1.1) DEFINITION. *For B a non-vacuous linear set the diameter of B is* $\sup B - \inf B$; *this is designated by the symbol* diam B. *For B vacuous* diam B *is defined as zero.*

It will presently be clear that diam B can be interpreted as the *essential diameter* of B without disturbing any of the subsequent results, where by essential diameter of B is meant ess sup $B -$ ess inf B when $|B|$ is > 0 and zero when $|B| = 0$.

Presented to the Society September 8, 1942; received by the editors August 22, 1942.

(1) Adams and Morse, *Random sampling in the evaluation of a Lebesgue integral*, Bull. Amer. Math. Soc. vol. 45 (1939) pp. 442–447, Theorem 4. Hereinafter this paper will be referred to as RS.

A set of real numbers will be spoken of as a *linear* set. For E a linear set measurable in the sense of Lebesgue and of positive measure, $L(E)$ represents the class of functions f the domain of each of which is E and for each of which the Lebesgue integral $\int_E f(x)dx$ exists (finite). For $f \in L(E)$ we usually abbreviate $\int_E f(x)dx$ to $\int_E f$.

When and only when a linear set E is measurable in the sense of Lebesgue its *measure* is designated by $|E|$.

All functions considered in this paper are *real-valued.*

(1.2) DEFINITION. *For E a measurable linear set and $0<\delta\leqq\infty$ a δ-partition F of E is a countable family of measurable, essentially disjoint sets satisfying the conditions*

$$\sum_{B\in F} B = E, \qquad \operatorname{diam} B < \delta \qquad \textit{for } B \in F.$$

F will be spoken of as an infinite family if and only if the sets $B\in F$ are infinite in number. The aggregate of all δ-partitions of E will be represented by $\Gamma_\delta(E)$.

(1.3) DEFINITION. *For E a measurable linear set $L^*(E)$ is the class of functions defined by the condition $f\in L^*(E)$ if and only if the domain of f is E and the condition*

$$-\infty < \int_B f < \infty$$

is satisfied for each measurable subset $B\subset E$ with

$$|B| > 0, \qquad \operatorname{diam} B < \infty.$$

Clearly we have $L(E)\subset L^*(E)$, with $L(E)=L^*(E)$ if and only if E is essentially bounded.

(1.4) DEFINITION. *For $f\in L^*(E)$ and B a measurable subset of E with $0<|B|<\infty$ we define*

$$\mathfrak{M}_B f = \int_B f/|B|.$$

The set of numbers of the form $\mathfrak{M}_B f$, where B is a measurable subset of E of positive measure with diam $B<\infty$, *will be designated by $\mathfrak{R}(f)$ and the closure of this set by $\overline{\mathfrak{R}(f)}$.*

It is easily seen that when $|E|$ is >0, $\mathfrak{R}(f)$ is an *interval*, and that this interval may be open, semi-open, or closed, as well as bounded or unbounded; the end points of this interval, whether finite or infinite, are ess $\inf_{x\in E} f(x)$ and ess $\sup_{x\in E} f(x)$. The subset of E on which f is not in $\mathfrak{R}(f)$ is of measure zero.

(1.5) DEFINITION. *For $f\in L^*(E)$ and ϕ a function whose domain includes $\mathfrak{R}(f)$ we interpret*

$$\phi[\mathfrak{M}_B f]\,|B| \qquad \textit{for } B\subset E,\ |B| = 0,$$

as zero. For $0<\delta\leqq\infty$ and $F\in\Gamma_\delta(E)$ we assign to the numerical sum

$$\sum_{B\in F}\phi[\mathfrak{M}_B f]\,|B|$$

its natural intuitional meaning (to be made precise later in Definition (2.7)) and

define two limits, which are actually of the nature of a lower integral and an upper integral respectively over E, as follows:

$$S_*(f, \phi, E) = \lim_{\delta\to 0} \inf_{F\in\Gamma_\delta(E)} \sum_{B\in F} \phi[\mathfrak{M}_B f] \,|B|,$$

$$S^*(f, \phi, E) = \lim_{\delta\to 0} \sup_{F\in\Gamma_\delta(E)} \sum_{B\in F} \phi[\mathfrak{M}_B f] \,|B|.$$

That each of these limits exists (finite or infinite) follows from the fact that

$$0 < \delta_1 \leqq \delta_2 \leqq \infty \quad \textit{implies} \quad \Gamma_{\delta_1}(E) \subset \Gamma_{\delta_2}(E).$$

When and only when these limits are the same we define

$$-\infty \leqq S(f, \phi, E) = S_*(f, \phi, E) = S^*(f, \phi, E) \leqq \infty.$$

Similarly we define two other limits, involving the sampling procedure, for any fixed real number $k \geqq 1 : S_*(f, \phi, E, k)$ $[S^*(f, \phi, E, k)]$ *is the* $\lim_{\delta\to 0}$ *of the* inf [sup] *of numbers of the form*

$$\sum_{B\in F} \phi[\mathfrak{M}_{AB} f]\,|B|,$$

where $F\in\Gamma_\delta(E)$ *and A is a measurable linear set variable with F and for each F variable within the restriction*

$$|B| \leqq k\,|AB| \qquad \textit{for } B \in F. \tag{1.6}$$

If and only if these limits are the same we write

$$-\infty \leqq S(f, \phi, E, k) = S_*(f, \phi, E, k) = S^*(f, \phi, E, k) \leqq \infty.$$

(1.7) Definition. *For E a linear set, f a function whose domain is E, and* ϕ *a function whose domain includes the range of f, we represent by* $\phi : f$ *the function g defined on E by the condition*

$$g(x) = \phi[f(x)] \qquad \textit{for } x \in E.$$

Our main problem is to determine conditions on f and ϕ which will insure the (finite) existence of $\int_E \phi : f$ and of $S(f, \phi, E, k)$ and the equality

$$\int_E \phi : f = S(f, \phi, E, k). \tag{1.8}$$

Actually the conditions which we obtain will insure

$$\begin{aligned} &\lim_{\delta\to 0} \sup_{F\in\Gamma_\delta(E)} \sum_{B\in F} \left| \int_B \phi : f - \phi[\mathfrak{M}_{AB} f]\,|B| \right| \\ &\quad = \lim_{\delta\to 0} \sup_{F\in\Gamma_\delta(E)} \sum_{B\in F} \left| \mathfrak{M}_B \phi : f - \phi[\mathfrak{M}_{AB} f] \right| |B| = 0, \end{aligned} \tag{1.9}$$

which implies (1.8).

For convenience of reference we attach numbers to the following particu-

lar cases of (1.8) and (1.9) in which no sampling is involved:

$$\int_E \phi:f = S(f, \phi, E), \tag{1.10}$$

$$\lim_{\delta\to 0} \sup_{F\in\Gamma_\delta(E)} \sum_{B\in F} \left| \mathfrak{M}_B\phi:f - \phi[\mathfrak{M}_B f] \right| |B| = 0. \tag{1.11}$$

The result quoted above from RS shows that the conditions $f\in L(E)$, $\phi(y)=y$ for $y\in\mathfrak{R}(f)$ are sufficient for (1.8). In RS the condition (1.9) could equally well have been obtained, but it was not. It has also been shown by Banach[2] that if $E=I$, the unit interval $0\leqq x\leqq 1$, and if the sets $B\in\Gamma_\delta(I)$ are restricted to be a *finite set of intervals*, the conditions $f\in L_p(I)$, $p>1$, $\phi(y)=|y|^p$ for $y\in\mathfrak{R}(f)$ are sufficient for (1.10).

It should be noted that if ϕ is regarded as defined on the infinite interval $-\infty<y<\infty$ and fixed, the subclass of $L(E)$ singled out by the condition $\int_E\phi:f$ exists may naturally be called $L_\phi(E)$, thus generalizing the familiar L_p classes of functions. This is by no means the first time that such generalizations have been considered. In fact they were introduced as early as 1924 by W. H. Young[3] who used the term "super-summability" in this connection. This notion was further developed, and related problems studied, by Young and by others including Burkill, Kaczmarz and Nikliborc, Orlicz, and Birnbaum and Orlicz[4]. There is no need to describe here these earlier investigations; it is sufficient to remark that the objectives of their authors were quite different from ours and were such as to require hypotheses on the function ϕ more restrictive than those which we shall impose.

2. **Further definitions and preliminaries.** We collect here a number of definitions which are to be used more or less generally throughout this paper. Other definitions, of only local use in the subsequent developments, will be formulated when needed.

(2.1) DEFINITION. *If Q is a condition involving the symbol a, $E_a[Q]$ stands for the set defined by the condition $b\in E_a[Q]$ if and only if Q is satisfied when b is substituted for a.*

(2.2) DEFINITION. *If A is a set and $a\in A$ we define $\{a\}$ as $E_x[x=a]$.*

(2) Banach, *Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales*, Fund. Math. vol. 3 (1922) pp. 133–181; especially pp. 175–177. We use $L_p(E)$ to stand for the class of functions f with $f\in L(E)$, $|f|^p\in L(E)$.

(3) W. H. Young, *The progress of mathematical analysis in the twentieth century*, Proc. London Math. Soc. (2) vol. 24 (1925–1926) pp. 421–434.

(4) Burkill, *The strong and weak convergence of functions of general type*, Proc. London Math. Soc. (2) vol. 28 (1928) pp. 493–500; Kaczmarz and Nikliborc, *Sur les suites des fonctions convergentes en moyenne*, Fund. Math. vol. 11 (1928) pp. 151–168; Orlicz, *Beiträge zur Theorie der Orthogonalentwicklungen*, Studia Mathematica vol. 1 (1929) pp. 1–39, 241–255; Birnbaum and Orlicz, *Über Approximation im Mittel*, ibid. vol. 2 (1930) pp. 197–206, and *Über die Verallgemeinerung des Begriffes der zueinander konjugierten Potenzen*, ibid. vol. 3 (1931) pp. 1–67.

(2.3) DEFINITION. *The symbol* $\sim$, *placed above another symbol, negates the meaning of the other symbol; in particular, if* A *is a set,* $a \tilde{\in} A$ *means that* a *is not an element of* A.

(2.4) DEFINITION. *For* a *and* b *real numbers with* $a \leqq b$ *we define* $[a, b]$ *as the closed interval* $E_x[a \leqq x \leqq b]$; *for* $a < b$ *the open interval* $E_x[a < x < b]$ *is represented by* (a, b). *The unit interval* $[0, 1]$ *is designated by* I.

(2.5) DEFINITION. *If* F *is a family of sets* B, *we sometimes employ for the set* $\sum_{B \in F} B$ *the abbreviation* $\sigma(F)$.

It should be noted that $x \in \sigma(F)$ if and only if there exists $B \in F$ with $x \in B$.

(2.6) DEFINITION. *Two sets* A *and* B *are called disjoint if and only if* AB *is vacuous (that is,* $AB = 0$). *A family* F *of sets* B *is called disjointed if and only if* $B_1 B_2 = 0$ *whenever* $B_1 \in F$, $B_2 \in F$ *and* $B_1 \neq B_2$.

Adopting the convention that $a + \infty = \infty$, $a - \infty = -\infty$ when $-\infty < a < \infty$ and that $\infty - \infty$ is meaningless, we set up the following formal definition for numerical sums.

(2.7) DEFINITION. *If* H *is a countable set with* N $(\leqq \infty)$ *elements and we have* $-\infty \leqq a_x \leqq \infty$ *for* $x \in H$, *the numerical sum* $\sum_{x \in H} a_x$ *is the unique number* λ $(-\infty \leqq \lambda \leqq \infty)$, *if any exists, such that*

$$\lambda = \sum_{j=1}^{N} a_{x_j}$$

whenever

$$H = \sum_{j=1}^{N} \{x_j\} \qquad \text{with } x_m \neq x_n \text{ for } 1 \leqq m < n < N + 1.$$

If H *is vacuous we understand*

$$\sum_{x \in H} a_x = 0.$$

It should be noted that $\left|\sum_{x \in H} a_x\right| < \infty$ implies $\sum_{x \in H} |a_x| < \infty$.

(2.8) DEFINITION. *For* J *a linear interval with* $0 < |J| \leqq \infty$ *a* δ*-partition of* J *in which each set* B *is an interval (open, semi-open, or closed) with* $|B| > 0$ *is called an interval* δ*-partition of* J; *the aggregate of all such interval* δ*-partitions is represented by* $\Gamma_\delta(J)$; *and limits corresponding to those of Definition* (1.5) *are designated by*

$$\overline{S}_*(f, \phi, J, k), \quad \overline{S}^*(f, \phi, J, k), \quad \textit{and so on.}$$

(2.9) DEFINITION. *For* E *a linear set the symbols* $C(E)$, $UC(E)$, *and* $B(E)$ *are used to designate, respectively, the classes of functions* f *the domain of each*

of which is E and each of which is continuous on E, uniformly continuous on E, and bounded on E. $R(I)$ stands for the class of functions each of which is (properly) Riemann integrable on I; $R^*(I)$ for the class of functions defined thus: $f\in R^*(I)$ if and only if for each $N>0$, $f_N\in R(I)$ where $f_N(x)=f(x)$ for $x\in E_x[-N\leqq f(x)\leqq N]$, $f_N(x)=-N$ for $x\in E_x[f(x)<-N]$, $f_N(x)=N$ for $x\in E_x[f(x)>N]$, and $\lim_{N\to\infty}\int_I f_N$ exists (finite). The notation $BC(E)$ [or $BUC(E)$ or $BL_\phi(E)$ or $\cdots$] indicates the intersection of the classes $B(E)$ and $C(E)$ [or $UC(E)$ or $L_\phi(E)$ or $\cdots$]. The condition $f\in L(I)-R(I)$ means that f is an element of $L(I)$ but not of $R(I)$. If A stands for any one of the classes mentioned here the relation $\phi\in A(E)$ is to be interpreted as meaning that ϕ is a function of the class A on E; in contrast, the relation $f\in A(E)$ [$\phi:f\in A(E)$] is understood to mean that f [$\phi:f$] is almost everywhere on E equal to a function of the class A on E.

In §3 we prove some theorems in the theory of functions of sets. These considerations are made in a generality considerably greater than is necessary for our particular purposes, since they seem to us to be fundamental to that theory and to possess some intrinsic interest. They are applied in §4, where we determine certain properties of the sums that appear in Definition (1.5).

In §5 we prove a sequence of theorems of which Theorem (5.12), the climax, gives a necessary and sufficient condition for (1.9) when $f\in L^*(E)$ and $\phi\in C(\Re(f))$. A sufficient condition, in terms of the existence of a convex dominant for $|\phi|$ and of such nature as frequently to be useful in testing, is given in Theorem (5.13). The necessity of this condition is the subject of inquiry in §§7, 8 following the establishment of a number of preliminary lemmas in §6.

Obviously the condition $f\in L^*(E)$ is necessary for the mere formulation of the main questions as expressed by (1.8), (1.9), (1.10), and (1.11). The object of §9 is to show that the hypothesis $\phi\in C(\Re(f))$ is also essential. In §10 we analyze the circumstances under which the k-hypothesis on the sample sets can be relaxed; that is, the condition (1.6) replaced by

$$|AB| > 0 \qquad \text{for } B\in F\in\Gamma_\delta(E). \tag{2.10}$$

In §11, we apply some of the earlier results to functions which are of bounded variation or are absolutely continuous in a certain generalized sense.

Our results hold also for a measurable set E in Euclidean n-space. Even in §6 where the treatment appears to be peculiarly 1-dimensional, one need only alter Definition (6.1) to employ n-dimensional measure in connection with the function u (and later f) and linear measure in connection with the equimeasurable function v (and later g).

In general definitions, theorems, and displayed relations to which other than local reference is made will be numbered sequentially in a decimal sys-

tem with the first number indicating a section. A displayed relation referred to only in the proof where it occurs will be labeled with a letter, such as (a).

3. **Some theorems in the theory of functions of sets.** In this section the symbol 0 will always stand for the vacuous set. We understand that two sets E_1 and E_2 are *identical*, and write $E_1 = E_2$, if and only if $x \in E_1$ implies and is implied by $x \in E_2$; otherwise we call the sets *distinct* and write $E_1 \neq E_2$. Under this definition of distinctness there is only one vacuous set. An aggregate of sets will usually be spoken of as a *family*.

(3.1) DEFINITION. *A family $\mathfrak{G}$ is called a σ-field if and only if* (i) *H a countable subfamily of $\mathfrak{G}$ implies $\sigma(H) \in \mathfrak{G}$, and* (ii) *$B \in \mathfrak{G}$ implies $\sigma(\mathfrak{G}) - B \in \mathfrak{G}$.*

If $\mathfrak{G}$ is an arbitrary family, $\beta \in \mathfrak{G}$ implies $\beta \subset \sigma(\mathfrak{G})$.

If $\mathfrak{G}$ is a σ-field, (i) implies $0 = \sigma(0) \in \mathfrak{G}$; from (ii) then follows $\sigma(\mathfrak{G}) \in \mathfrak{G}$; whence we find that $B_1 \in \mathfrak{G}$, $B_2 \in \mathfrak{G}$ imply $B_1B_2 \in \mathfrak{G}$ and $B_2 - B_1 = B_2 - B_1B_2 \in \mathfrak{G}$.

The simplest and "smallest" σ-field is $\{0\}$. If S represents Euclidean n-space ($1 \leq n < \infty$) and $E \subset S$ is a measurable set (of finite or infinite measure in the sense of Lebesgue), the measurable subsets of E constitute a σ-field $\mathfrak{G}$ with $\sigma(\mathfrak{G}) = E$.

(3.2) DEFINITION. *If A and B are sets, we define*

$$A * B = A \mathop{E}_{\beta} [\beta \subset B].$$

The following theorem may easily be verified.

(3.3) THEOREM. *If $\mathfrak{G}$ is a σ-field and $B \in \mathfrak{G}$, then $\mathfrak{G} * B$ is a σ-field with $\sigma(\mathfrak{G} * B) = B$.*

(3.4) DEFINITION. *For $\mathfrak{G}$ a σ-field F is called a $\mathfrak{G}$-partition if and only if F is a countable disjointed subfamily of $\mathfrak{G}$ with $\sigma(F) = \sigma(\mathfrak{G})$. The aggregate of all $\mathfrak{G}$-partitions is represented by $\mathfrak{G}^\pi$.*

It may be noted that if $\mathfrak{G}$ is a σ-field, $\{\sigma(\mathfrak{G})\} \in \mathfrak{G}^\pi$; thus $\mathfrak{G}^\pi$ is never vacuous. Also $F \in \mathfrak{G}^\pi$ implies $F - \{0\} \in \mathfrak{G}^\pi$.

(3.5) DEFINITION. *If $\mathfrak{G}$ is a σ-field and f a function whose domain includes $\mathfrak{G}$, f is said to be quasi-additive on $\mathfrak{G}$ when and only when*

$$\sup_{F \in \mathfrak{G}^\pi} \left| \sum_{B \in F} f(B) \right| < \infty.$$

A well known theorem on the summation of absolutely convergent double series yields a result which we find convenient for our present purposes to put in the following form.

(3.6) THEOREM. *If F is a countable family and for $B \in F$, K_B is a countable*

family; $K_{B_1}K_{B_2}=0$ *whenever* $B_1\in F$, $B_2\in F$, $B_1\neq B_2$; $S=\sum_{B\in F}K_B$; *and* f *is a function whose domain includes* S *and which satisfies the condition* $|\sum_{\beta\in S}f(\beta)| < \infty$ (*and therefore* $\sum_{\beta\in S}|f(\beta)| < \infty$), *then*

$$\sum_{\beta\in S} f(\beta) = \sum_{B\in F}\ \sum_{\beta\in K_B} f(\beta).$$

In the rest of this section frequent use will be made of this theorem.

(3.7) THEOREM. *If* f *is quasi-additive on* $\mathfrak{G}$ *and* $B\in\mathfrak{G}$, *then* f *is quasi-additive on* $\mathfrak{G}*B$.

Proof. Let $E=\sigma(\mathfrak{G})$. In case $E-B=0$ we have $\mathfrak{G}*B=\mathfrak{G}*E=\mathfrak{G}*\sigma(\mathfrak{G})=\mathfrak{G}$ and the theorem is trivially true. If $E-B\neq 0$, we let F be an arbitrary element of $(\mathfrak{G}*B)^\pi$, set $F'=\{E-B\}+F$, note that $F'\in\mathfrak{G}^\pi$, and observe that the relation

$$\left|\sum_{\beta\in F} f(\beta)\right| \leqq |-f(E-B)| + \left|f(E-B)+\sum_{\beta\in F} f(\beta)\right|$$

$$= |f(E-B)| + \left|\sum_{\beta\in F'} f(\beta)\right| \leqq |f(E-B)| + \sup_{F''\in\mathfrak{G}^\pi}\left|\sum_{\beta\in F''} f(\beta)\right|$$

completes the proof.

(3.8) LEMMA. *If* S *is quasi-additive on* $\mathfrak{G}$, $A\in\mathfrak{G}$, $B\in\mathfrak{G}$, $A+B=\sigma(\mathfrak{G})$, $AB=0$, *and* f *is bounded on* $\mathfrak{G}*A$ *and on* $\mathfrak{G}*B$, *then* f *is bounded on* $\mathfrak{G}$.

Proof. Let $E=\sigma(\mathfrak{G})$,

$$N=\sup_{F\in\mathfrak{G}^\pi}\left|\sum_{\beta\in F} f(\beta)\right|,\qquad N_A=\sup_{\beta\in\mathfrak{G}*A}|f(\beta)|,\qquad N_B=\sup_{\beta\in\mathfrak{G}*B}|f(\beta)|.$$

Let γ' be an arbitrary element of $\mathfrak{G}$, $\alpha'=A(E-\gamma')$, $\beta'=B(E-\gamma')$, and note that $\alpha'\beta'=\beta'\gamma'=\gamma'\alpha'=0$. If $\alpha'+\gamma'=0$, the relation

$$|f(\gamma')| \leqq N+|f(\alpha')|+|f(\beta')| \leqq N+N_A+N_B$$

is apparent; if $\alpha'+\gamma'\neq 0$ and $\beta'=0$, it is a consequence of the inequality

$$|f(\alpha')+f(\gamma')| \leqq N;$$

and if $\alpha'+\gamma'\neq 0$ and $\beta'\neq 0$, it follows from the inequality

$$|f(\alpha')+f(\beta')+f(\gamma')| \leqq N.$$

Thus the proof is complete.

(3.9) LEMMA. *If* f *is quasi-additive but unbounded on* $\mathfrak{G}$, *there exist sets* A *and* B *with*

$$A\in\mathfrak{G},\quad B\in\mathfrak{G},\quad A+B=\sigma(\mathfrak{G}),\quad AB=0,\quad |f(A)|>1+|f(0)|,$$

f unbounded on $\mathfrak{S} * B$.

Proof. Let E and N have the meaning assigned them in the last proof. Since f is unbounded on $\mathfrak{S}$ we take $B_0 \in \mathfrak{S}$ such that

$$|f(B_0)| > N + 1 + |f(0)|,$$

and let $A_0 = E - B_0$. Clearly we have $B_0 \neq 0$ and

$$A_0 \in \mathfrak{S}, \quad B_0 \in \mathfrak{S}, \quad A_0 + B_0 = E, \quad A_0B_0 = 0, \quad |f(B_0)| > 1 + |f(0)|,$$

$$\begin{aligned} |f(A_0)| &\geqq |f(B_0)| - |f(A_0) + f(B_0)| \\ &> N + 1 + |f(0)| - N = 1 + |f(0)|. \end{aligned}$$

From Lemma (3.8) we infer that f is either unbounded on $\mathfrak{S} * B_0$ or unbounded on $\mathfrak{S} * A_0$; in the first alternative the desired conclusion is reached by taking $A = A_0$ and $B = B_0$, whereas in the second case it is obtained by taking $A = B_0$ and $B = A_0$.

A proof of a well known theorem on completely additive functions of sets(6), which was communicated to us some time ago by R. M. Robinson, has been of use to us at this juncture. The following is a generalization of that theorem.

(3.10) THEOREM. *If f is quasi-additive on $\mathfrak{S}$, f is bounded on $\mathfrak{S}$.*

Proof. We shall show that the assumption of the hypothesis and the contrary of the conclusion leads to a contradiction. Let $B_0 = \sigma(\mathfrak{S})$. In view of Theorem (3.7), Lemma (3.9), and Theorem (3.3) there exist sets $A_1, A_2, A_3, \cdots$ and $B_1, B_2, B_3, \cdots$ satisfying the following recursive condition: if n is a positive integer and

$$A_n \in \mathfrak{S} * B_{n-1}, \quad B_n \in \mathfrak{S} * B_{n-1}, \quad A_n + B_n = B_{n-1}, \quad A_nB_n = 0,$$
$$|f(A_n)| > 1 + |f(0)|, \qquad f \text{ is unbounded on } \mathfrak{S} * B_n,$$

then we have

$$A_{n+1} \in \mathfrak{S} * B_n, \quad B_{n+1} \in \mathfrak{S} * B_n, \quad A_{n+1} + B_{n+1} = B_n, \quad A_{n+1}B_{n+1} = 0,$$
$$|f(A_{n+1})| > 1 + |f(0)|, \qquad f \text{ is unbounded on } \mathfrak{S} * B_{n+1}.$$

Clearly we also have

$$B_0 \supset B_1 \supset B_2 \supset \cdots,$$
$$A_mB_n \subset A_mB_{n-1} \subset A_mB_m = 0 \qquad \text{for } 1 \leqq m < n.$$

Defining

$$A_0 = B_0 - \sum_{j=1}^{\infty} A_j, \qquad F_0 = \sum_{j=0}^{\infty} \{A_j\},$$

we see that $A_mA_n = 0$ for $0 \leqq m < n$. From this and the fact that

(6) See, for example, Saks, *Theory of the integral*, Warsaw, 1937, Theorem (6.1), p. 10.

$$|f(A_n)| > 1 + |f(0)| \quad \text{implies} \quad A_n \neq 0 \qquad (n = 1, 2, 3, \cdots),$$

we infer

$$A_m \neq A_n \qquad \text{for } 0 \leqq m < n.$$

But F_0 is a $\mathfrak{G}$-partition; hence

$$\left|\sum_{n=0}^{\infty} f(A_n)\right| = \left|\sum_{\beta\in F_0} f(\beta)\right| \leqq N < \infty$$

and $f(A_n)$ tends to zero as $n\to\infty$, in contradiction to the inequalities $|f(A_n)| > 1+|f(0)|$ $(n=1, 2, 3, \cdots)$.

(3.11) Definition. *A family H is called hereditary if and only if the relation $B\subset A\in H$ implies $B\in H$.*

It may be noted that the families 0 and $\{0\}$ are hereditary.

For $0<\delta\leqq\infty$ let H_δ represent the family of all subsets of the real number system having diameter $<\delta$; H_δ is then hereditary. If E is a measurable set of real numbers and $\mathfrak{G}$ represents the family of all measurable subsets of E, $\mathfrak{G}$ is a σ-field and we have

$$\mathfrak{G}^{\tau}*H_\delta = \Gamma_\delta(E).$$

(3.12) Theorem. *If $\mathfrak{G}$ is a σ-field, H is an hereditary family, and f is a function whose domain includes $\mathfrak{G}H$, then the relation*

$$\sup_{F\in\mathfrak{G}^{\tau}*H} \left|\sum_{\beta\in F} f(\beta)\right| = N < \infty$$

implies

$$\sup_{F\in\mathfrak{G}^{\tau}*H} \sum_{\beta\in F} |f(\beta)| < \infty.$$

Proof. Since $\sup 0 = -\infty$, the theorem is trivial in the case of $\mathfrak{G}^{\tau}*H=0$. Henceforth we assume $\mathfrak{G}^{\tau}*H\neq 0$; let $E=\sigma(\mathfrak{G})$; let V be some fixed element of $\mathfrak{G}^{\tau}*H$; for $\beta\in\mathfrak{G}$ let

$$V_\beta = V \mathop{E}_{B} [B\beta \neq 0], \qquad K_\beta = \sum_{B\in V_\beta} \{B\beta\};$$

and let g be a function with

$$g(\beta) = \sum_{B\in V_\beta} f(B\beta) \qquad \text{for } \beta\in\mathfrak{G}.$$

The first step in the proof is to show that the function g is quasi-additive on $\mathfrak{G}$. For any $F\in\mathfrak{G}^{\tau}$ let

$$W = \sum_{\beta\in F} K_\beta.$$

W is then an element of $\mathfrak{S}^r$ and since H is hereditary we infer $W \in \mathfrak{S}^r \cdot H$. Moreover, it is readily seen that $K_{\beta_1} K_{\beta_2} = 0$ whenever $\beta_1 \in F$, $\beta_2 \in F$, $\beta_1 \neq \beta_2$ and also that $\{\beta_1 \beta\}\{\beta_2 \beta\} = 0$ whenever $\beta \in F$, $\beta_1 \in V_\beta$, $\beta_2 \in V_\beta$, $\beta_1 \neq \beta_2$. Two applications of Theorem (3.7) thus yield the relation

$$\sum_{\gamma \in W} f(\gamma) = \sum_{\beta \in F} \sum_{\gamma \in K_\beta} f(\gamma) = \sum_{\beta \in F} \sum_{B \in V_\beta} \sum_{\gamma \in \{B\beta\}} f(\gamma)$$
$$= \sum_{\beta \in F} \sum_{B \in V_\beta} f(B\beta) = \sum_{\beta \in F} g(\beta),$$

an immediate implication of which is the inequality

$$\left| \sum_{\beta \in F} g(\beta) \right| \leqq N.$$

The second step is to establish the existence of a number M with

$$\left| \sum_{\beta \in D} f(\beta) \right| \leqq M < \infty$$

for every countable disjointed subfamily D of $\mathfrak{S}H$. Using the result of the first step and Theorem (3.10) we infer the existence of a number N_1 with

$$| g(\beta) | \leqq N_1 < \infty \qquad \text{for } \beta \in \mathfrak{S}.$$

Let $M = N + N_1$; let D be any countable disjointed subfamily of $\mathfrak{S}H$; let $C = E - \sigma(D)$; let

$$F_1 = D + \sum_{B \in V_C} \{BC\},$$

and note that $F_1 \in \mathfrak{S}^r \cdot H$. We have

$$\sum_{\beta \in F_1} f(\beta) = \sum_{\beta \in D} f(\beta) + \sum_{B \in V_C} \sum_{\beta \in \{BC\}} f(\beta) = \sum_{\beta \in D} f(\beta) + \sum_{B \in V_C} f(BC)$$
$$= \sum_{\beta \in D} f(\beta) + g(C),$$

whence

$$\left| \sum_{\beta \in D} f(\beta) \right| \leqq N + N_1 = M.$$

Finally, let

$$F \in \mathfrak{S}^r \cdot H, \qquad D_1 = F \mathop{E}_{\beta} [f(\beta) \geqq 0], \qquad D_2 = F \mathop{E}_{\beta} [f(\beta) < 0].$$

We then have the desired conclusion:

$$\sum_{\beta \in F} | f(\beta) | = \left| \sum_{\beta \in D_1} f(\beta) \right| + \left| \sum_{\beta \in D_2} f(\beta) \right| \leqq 2M < \infty.$$

It may be remarked that if in Theorem (3.12) we let H stand for the hereditary family $E_C[C\subset\sigma(\mathfrak{S})]$, then $\mathfrak{S}^\nu * H=\mathfrak{S}^\nu$.

4. **Concerning the behavior of the sums.** The sums considered in Definition (1.5) have a considerable number of properties which are seen immediately and which it is essential to note. Partly for convenience of reference we assemble some of these properties in

(4.1) THEOREM. *Let $f\in L^*(E)$ and ϕ and ϕ_1 be functions each of whose domains includes $\mathfrak{R}(f)$.*

(i) *Since $0<\delta_1\leqq\delta_2\leqq\infty$ implies $\Gamma_{\delta_1}(E)\subset\Gamma_{\delta_2}(E)$, the* $\inf_{F\in\Gamma_\delta(E)}$ $[\sup_{F\in\Gamma_\delta(E)}]$ *of sums of the form*

$$\sum_{B\in F}\phi[\mathfrak{M}_{AB}f]\,|B| \quad \textit{or} \quad \sum_{B\in F}\left|\mathfrak{M}_B\phi:f-\phi[\mathfrak{M}_{AB}f]\right|\,|B|$$

is non-decreasing [non-increasing] as δ decreases ($0<\delta\leqq\infty$).

For $0<\delta\leqq\infty$ and $F\in\Gamma_\delta(E)$ we have the following properties of sums:

(ii) *$k=1$ implies $|AB|=|B|$ for $B\in F$, $|AE|=|E|$, and*

$$\sum_{B\in F}\phi[\mathfrak{M}_{AB}f]\,|B|=\sum_{B\in F}\phi[\mathfrak{M}_Bf]\,|B|;$$

(iii) *$0\leqq\phi(y)$ for $y\in\mathfrak{R}(f)$ implies*

$$\sum_{B\in F}\phi[\mathfrak{M}_{AB}f]\,|B|\leqq k\sum_{B\in F}\phi[\mathfrak{M}_{AB}f]\,|AB|$$

and

$$\sup_{F\in\Gamma_\delta(E)}\sum_{B\in F}\phi[\mathfrak{M}_{AB}f]\,|B|\leqq k\sup_{F\in\Gamma_\delta(E)}\sum_{B\in F}\phi[\mathfrak{M}_Bf]\,|B|;$$

(iv) *$\phi(y)\leqq\phi_1(y)$ for $y\in\mathfrak{R}(f)$ implies*

$$\sum_{B\in F}\phi[\mathfrak{M}_{AB}f]\,|B|\leqq\sum_{B\in F}\phi_1[\mathfrak{M}_{AB}f]\,|B|;$$

(v) *$(\phi+\phi_1)(y)=\phi(y)+\phi_1(y)$ for $y\in\mathfrak{R}(f)$ implies*

$$\sum_{B\in F}(\phi+\phi_1)[\mathfrak{M}_{AB}f]\,|B|=\sum_{B\in F}\phi[\mathfrak{M}_{AB}f]\,|B|+\sum_{B\in F}\phi_1[\mathfrak{M}_{AB}f]\,|B|$$

and

$$\begin{aligned}&\inf_{F\in\Gamma_\delta(E)}\sum_{B\in F}\phi[\mathfrak{M}_{AB}f]\,|B|+\inf_{F\in\Gamma_\delta(E)}\sum_{B\in F}\phi_1[\mathfrak{M}_{AB}f]\,|B|\\ &\quad\leqq\inf_{F\in\Gamma_\delta(E)}\sum_{B\in F}(\phi+\phi_1)[\mathfrak{M}_{AB}f]\,|B|\leqq\sup_{F\in\Gamma_\delta(E)}\sum_{B\in F}(\phi+\phi_1)[\mathfrak{M}_{AB}f]\,|B|\\ &\quad\leqq\sup_{F\in\Gamma_\delta(E)}\sum_{B\in F}\phi[\mathfrak{M}_{AB}f]\,|B|+\sup_{F\in\Gamma_\delta(E)}\sum_{B\in F}\phi_1[\mathfrak{M}_{AB}f]\,|B|.\end{aligned}$$

Corresponding properties of limits of sums are:

(vi) $S_*(f,\phi,E,1)=S_*(f,\phi,E)\leqq S^*(f,\phi,E)=S^*(f,\phi,E,1)$;

(vii) $0 \leqq \phi(y)$ *for* $y \in \Re(f)$ *implies*

$$S^*(f, \phi, E, k) \leqq kS^*(f, \phi, E);$$

(viii) $\phi(y) \leqq \phi_1(y)$ *for* $y \in \Re(f)$ *implies*

$$S_*(f, \phi, E, k) \leqq S_*(f, \phi_1, E, k), \qquad S^*(f, \phi, E, k) \leqq S^*(f, \phi_1, E, k);$$

(ix) $(\phi+\phi_1)(y) = \phi(y) + \phi_1(y)$ *for* $y \in \Re(f)$ *implies*

$$S_*(f, \phi, E, k) + S_*(f, \phi_1, E, k) \leqq S_*(f, \phi + \phi_1, E, k)$$
$$\leqq S^*(f, \phi + \phi_1, E, k) \leqq S^*(f, \phi, E, k) + S^*(f, \phi_1, E, k).$$

It must be emphasized that when we write, as in the theorem above,

$$\sum_{B \in F} \phi[\mathfrak{M}_{AB} f] \mid B \mid,$$

without inf or sup ahead of it, A is understood to be an arbitrary *fixed* set satisfying condition (1.6); whereas, when

$$\inf_{F \in \Gamma_\delta(E)} \quad \text{or} \quad \sup_{F \in \Gamma_\delta(E)} \sum_{B \in F} \phi[\mathfrak{M}_{AB} f] \mid B \mid$$

is written, A is understood to *vary* with F with the freedom allowed by condition (1.6). To define a notation that makes this distinction would introduce, it seems to us, an unnecessary complication.

The next four theorems are fundamental for later developments.

(4.2) THEOREM. *If* $f \in L^*(E)$, ϕ *is a function whose domain includes* $\Re(f)$, *and* $0 < \delta \leqq \infty$, *the condition*

$$\sup_{F \in \Gamma_\delta(E)} \left| \sum_{B \in F} \phi[\mathfrak{M}_B f] \mid B \mid \right| < \infty$$

implies

$$\sup_{F \in \Gamma_\delta(E)} \sum_{B \in F} \mid \phi[\mathfrak{M}_B f] \mid \mid B \mid < \infty.$$

Taking into account the remark immediately preceding the statement of Theorem (3.12), one sees at once that this theorem is a corollary of Theorem (3.12).

COROLLARY. *If* $f \in L^*(E)$ *and* ϕ *is a function whose domain includes* $\Re(f)$, *the condition*

$$-\infty < S_*(f, \phi, E) \leqq S^*(f, \phi, E) < \infty$$

implies $0 \leqq S^*(f, |\phi|, E) < \infty$.

(4.3) THEOREM. *If* $f \in L^*(E)$, $0 \leqq \phi(y)$ *for* $y \in \Re(f)$, *and* $0 < \delta \leqq \infty$, *the condition*

$$\sup_{F\in\Gamma_\delta(E)} \sum_{B\in F} \phi[\mathfrak{M}_B f]\,|B| = M < \infty$$

implies the existence of $N=N(\epsilon)>0$ *with*

$$0 < N < \infty, \quad \sup_{F\in\Gamma_\delta(H)} \sum_{B\in F} \phi[\mathfrak{M}_B f]\,|B| < \epsilon \quad \textit{for} \quad H \subset E - [-N, N].$$

Proof. Let $F_1\in\Gamma_\delta(E)$ be such that

$$\sum_{B\in F_1} \phi[\mathfrak{M}_B f]\,|B| > M - \epsilon.$$

Clearly F_1 contains a finite family F_2 satisfying the same inequality. Let N satisfy the conditions

$$0 < N < \infty, \qquad \sigma(F_2) \subset [-N, N].$$

For $H\subset E-[-N, N]$ and $F\in\Gamma_\delta(H)$, we then have the desired condition.

(4.4) Theorem. *If* $f\in L^*(E)$, $\phi\in C(\mathfrak{R}(f))$, $0\leqq\phi(y)$ *for* $y\in\mathfrak{R}(f)$, $0<\delta\leqq\infty$,

$$\sup_{F\in\Gamma_\delta(H)} \sum_{B\in F} \phi[\mathfrak{M}_B f]\,|B| < \infty,$$

and H *is a measurable subset of* E, *then we have*

$$\lim_{|H|\to 0}\ \sup_{F\in\Gamma_\delta(H)} \sum_{B\in F} \phi[\mathfrak{M}_B f]\,|B| = 0.$$

Proof. Let

$$\Lambda(B) = \phi[\mathfrak{M}_B f]\,|B| \qquad \text{for } B \in F \in \Gamma_\delta(H),$$
$$M = \limsup_{|H|\to 0} \sum_{B\in F\in\Gamma_\delta(H)} \Lambda(B),$$

M being finite on account of the hypotheses. If M is >0, there exists $\eta>0$ such that $|H|<\eta$ implies

$$\sum_{B\in F\in\Gamma_\delta(H)} \Lambda(B) \leqq 3M/2.$$

Let H_1 be a particular set H and $F_1\in\Gamma_\delta(H_1)$ a particular family of sets satisfying the conditions

$$|H_1| < \eta/2, \qquad \sum_{B\in F_1} \Lambda(B) > 4M/5.$$

It is clear that F_1 can be taken as a finite family. If D is a measurable subset of E, we have

$$\lim_{|D|\to 0} \int_{B-D} f/|B-D| = \int_B f/|B| \qquad \text{for } B \in F_1,\ |B| > 0;$$

hence, in view of the continuity of ϕ on $\mathfrak{R}(f)$ and the finiteness of the family

F_1, there exists γ with $0<\gamma<\eta/2$ such that $|D|<\gamma$ implies

$$\sum_{B\in F_1} \Lambda(B-D) > 4M/5.$$

Let H_2 be a particular set H and $F_2\in\Gamma_\delta(H_2)$ a particular family of sets satisfying the conditions

$$|H_2|<\gamma, \qquad \sum_{B\in F_2}\Lambda(B) > 4M/5.$$

Let

$$G = \sum_{B\in F_1}\{B-H_2\}$$

and consider the family $G+F_2$ of disjoint sets. We have

$$\sum_{B\in G+F_2}|B| \leqq \sum_{B\in F_1}|B| + \sum_{B\in F_2}|B| = |H_1| + |H_2| < \eta/2+\gamma<\eta,$$

whence

$$3M/2 \geqq \sum_{B\in G+F_2}\Lambda(B) = \sum_{B\in G}\Lambda(B) + \sum_{B\in F_2}\Lambda(B) > 4M/5 + 4M/5,$$

a contradiction which implies $M=0$ and completes the proof.

Combining Theorems (4.2) and (4.4) we obtain the following result.

(4.5) THEOREM. *If $f\in L^*(E)$, $\phi\in C(\mathfrak{R}(f))$, and H is a measurable subset of E, the condition*

$$-\infty < S_*(f,\phi,E) \leqq S^*(f,\phi,E) < \infty$$

implies

$$\lim_{|H|\to 0} S^*(f,|\phi|,H) = 0.$$

In other words, the conclusion of this theorem is that $S^*(f,|\phi|,H)$ is an absolutely continuous function of a measurable set $H\subset E$.

5. **Sufficient conditions for** (1.9). Since (1.9) is trivially true when $|E|=0$, we are justified in assuming $|E|>0$ in connection with the proof of each theorem in this section. The first theorem is an extension of the lemma of RS.

(5.1) THEOREM. *If E is essentially bounded, the conditions $f\in UC(E)$ and $\phi\in C(\overline{\mathfrak{R}(f)})$ are sufficient for* (1.9), *even with* (1.6) *replaced by* (2.10).

Proof. We may assume f itself to be uniformly continuous on E; then f is bounded on E, and $\overline{\mathfrak{R}(f)}$ is seen to be a bounded closed interval. Thus $\phi\in C(\overline{\mathfrak{R}(f)})$ implies $\phi\in UC(\overline{\mathfrak{R}(f)})$; $\phi:f$ is uniformly continuous on E; and we have

$$\int_B \phi:f = \sum_{B\in F}\int_B \phi:f \qquad \text{for } F\in\Gamma_\delta(E),\ 0<\delta\leqq\infty.$$

Let $\eta=\eta(\phi, \epsilon)>0$ be such that

$$y_1, y_2 \in \overline{\mathfrak{R}(f)}, \quad |y_1 - y_2| \leqq \eta \quad \text{imply} \quad |\phi(y_1) - \phi(y_2)| \leqq \epsilon/(2|E|);$$

and let $\delta=\delta(f, \eta)>0$ be such that

$$x_1, x_2 \in E, \quad |x_1 - x_2| \leqq \delta \quad \text{imply} \quad |f(x_1) - f(x_2)| \leqq \eta.$$

For $F\in\Gamma_\delta(E)$, $B\in F$, $|B|>0$, and $x_1 \in B$ we then have

$$\alpha = \inf_{x\in B} \phi[f(x)] \leqq \mathfrak{M}_B\phi:f, \ \phi[f(x_1)] \leqq \sup_{x\in B} \phi[f(x)] = \beta, \quad \beta - \alpha \leqq \epsilon/(2|E|),$$

$$a = \inf_{x\in B} f(x) \leqq \mathfrak{M}_{AB}f, \qquad f(x_1) \leqq \sup_{x\in B} f(x) = b, \qquad b - a \leqq \eta,$$

whence

$$\text{(a)} \qquad \begin{aligned} |\mathfrak{M}_B\phi:f - \phi[\mathfrak{M}_{AB}f]| &\leqq |\mathfrak{M}_B\phi:f - \phi[f(x_1)]| \\ &\quad + |\phi[f(x_1)] - \phi[\mathfrak{M}_{AB}f]| \leqq \epsilon/|E|, \end{aligned}$$

and

$$\sum_{B\in F} |\mathfrak{M}_B\phi:f - \phi[\mathfrak{M}_{AB}f]|\,|B| \leqq \epsilon.$$

REMARK. The relation (1.9) can be established by means of inequality (a) on the individual terms of (1.9) only under very special circumstances, such as those hypothesized in Theorem (5.1). That this cannot be done even in the simple case of $E=I$, $\phi(y)=y$ for $y\in\overline{\mathfrak{R}(f)}$, when k is >1 and f has a single point of essential discontinuity in I is illustrated by taking f on I to be the characteristic function of the set $E_x[1/2\leqq x\leqq 1]$.

(5.2) THEOREM. *If E is essentially bounded, the conditions $f\in L^*(E)=L(E)$ and $\phi\in UC(\mathfrak{R}(f))$ are sufficient for* (1.9).

Proof. If the interval $\mathfrak{R}(f)$ is not a closed set, it has a finite endpoint which does not belong to it. The hypothesis $\phi\in UC(\mathfrak{R}(f))$ implies that $\lim \phi(y)$ exists as $y\in\mathfrak{R}(f)$ approaches this endpoint. If at each such endpoint ϕ is defined (or redefined) as this limit, we have $\phi\in UC(\overline{\mathfrak{R}(f)})$. Clearly this extension of the definition of ϕ (or this redefinition of ϕ) does not affect the sums in question; neither does it affect the value of $\int_E \phi:f$, since the set of points $E\ E_x[f(x)\,\bar{\in}\,\mathfrak{R}(f)]$ is of measure zero. In other words, we may assume $\phi\in UC(\overline{\mathfrak{R}(f)})$.

Let $0<\epsilon_1<\epsilon/(1+|E|)$ and $\eta=\eta(\phi, \epsilon_1)>0$ be such that

$$y_1, y_2 \in \overline{\mathfrak{R}(f)}, \quad |y_1 - y_2| \leqq \eta \quad \text{imply} \quad |\phi(y)_1 - \phi(y_2)| \leqq \epsilon_1/4;$$

then we have

$$|\phi(y_1) - \phi(y_2)| \leqq \epsilon_1/4 + \epsilon_1|y_1 - y_2|/(4\eta) \quad \text{for} \quad y_1, y_2 \in \overline{\mathfrak{R}(f)}.$$

Let $K=[a, b]$ be a bounded closed interval containing all of E except for

a set of measure zero which can and will be neglected; let f be a particular function summable on E; and let x_1 satisfy the conditions

$$x_1 \in E, \quad \operatorname*{ess\,inf}_{x\in E} f(x) \leqq f(x_1) \leqq \operatorname*{ess\,sup}_{x\in E} f(x), \qquad -\infty < f(x_1) < \infty.$$

Then f can be defined (or redefined) as $f(x_1)$ for $x\in K-E$, so that f is summable on $[a, b]$ and the new $\Re(f)$ is the same as the old $\Re(f)$. On K the function f can be approximated arbitrarily closely in the mean by a function g continuous on K and having the property

$$\operatorname*{ess\,inf}_{x\in K} f(x) \leqq \operatorname*{ess\,inf}_{x\in K} g(x) \leqq \operatorname*{ess\,sup}_{x\in K} g(x) \leqq \operatorname*{ess\,sup}_{x\in K} f(x).$$

Thus we are assured of the existence of a function g satisfying the conditions

$$g \in UC(E), \qquad \int_E |f-g| < \eta/k \leqq \eta, \qquad \overline{\Re(g)} \subset \overline{\Re(f)}.$$

In accordance with Theorem (5.1) let $\delta = \delta\ (g, \phi, \epsilon_1/4) > 0$ correspond to $\epsilon_1/4$ for the functions g and ϕ. The function $\phi:f$ is measurable on E and we infer that it is summable on E from the inequality

$$|\phi[f(x)]| \leqq |\phi[g(x)]| + \epsilon_1/4 + \epsilon_1 |f(x) - g(x)| /(4\eta) \qquad \text{for } x \in E,$$

which expresses the dominance of $|\phi:f|$ by a summable function. For $F\in\Gamma_\delta(E)$ we now have

$$\begin{aligned}
&\sum_{B\in F}\left|\int_B \phi:f - \phi[\mathfrak{M}_{AB}f]\,|B|\right| \\
&\leqq \sum_{B\in F}\Bigg(\left|\int_B \phi:f - \int_B \phi:g\right| + \left|\int_B \phi:g - \phi[\mathfrak{M}_{AB}g]\,|B|\right| \\
&\qquad\qquad + |\phi[\mathfrak{M}_{AB}g] - \phi[\mathfrak{M}_{AB}f]|\,|B|\Bigg) \\
&\leqq \int_E |\phi:f - \phi:g| + \epsilon_1/4 + \sum_{B\in F}\left(\epsilon_1/4 + \epsilon_1 \int_{AB} |g-f|/(4\eta|AB|)\right)|B| \\
&\leqq \epsilon_1/4 + \epsilon_1 \int_E |f-g|/(4\eta) + \epsilon_1/4 + (\epsilon_1/4)|E| \\
&\qquad\qquad + k\epsilon_1 \int_E |g-f|/(4\eta) < \epsilon_1 + \epsilon_1|E| < \epsilon.
\end{aligned}$$

For use in proving subsequent theorems we introduce two definitions and a lemma.

(5.3) DEFINITION. *If $f\in L^*(E)$, $\sum_{i=1}^{\infty} J_i$ will be called a normal expression for $\Re(f)$ if and only if the following conditions are satisfied*: (1) *for each i, J_i is a*

bounded closed interval $[a_i, b_i] \subset \mathfrak{R}(f)$, *with*

$$[a_i, b_i] \subset J_i \subset J_{i+1} \qquad (i = 1, 2, 3, \cdots);$$

(2) *if* $\inf \mathfrak{R}(f) = a$ *is finite and* $a \in \mathfrak{R}(f)$, $a \in J_i$ $(i=1, 2, 3, \cdots)$; (3) *if* $\sup \mathfrak{R}(f) = b$ *is finite and* $b \in \mathfrak{R}(f)$, $b \in J_i$ $(i=1, 2, 3, \cdots)$; *and* (4) $\sum_{i=1}^{\infty} J_i = \mathfrak{R}(f)$.

It should be observed that if $\mathfrak{R}(f)$ is a bounded closed interval, $J_i = \mathfrak{R}(f)$ for each i.

(5.4) LEMMA. *Let* $f \in L(E)$ *and* $\sum_{i=1}^{\infty} J_i$ *be a normal expression for* $\mathfrak{R}(f)$. *For* $0 < \delta \leqq \infty$ *and* $F \in \Gamma_\delta(E)$ *let* $G_i \subset F$ *represent the family of sets* $B \in F$ *defined by the condition*

$$\mathfrak{M}_B f \in \mathfrak{R}(f) - J_i \qquad (i = 1, 2, 3, \cdots).$$

Then we have

$$\lim_{i \to \infty} \sup_{F \in \Gamma_\delta(E)} \sum_{B \in G_i} |B| = 0.$$

Proof. In case $\mathfrak{R}(f)$ is a bounded closed interval, $\mathfrak{R}(f) - J_i$ is vacuous and $\sum_{B \in G_i} |B| = 0$ for each i. In case $\sup \mathfrak{R}(f) = \infty$ we have $\lim_{i \to \infty} b_i = \infty$. For each i such that $b_i > 0$, let $H_i \subset F$ represent the family of sets B defined by the condition $\mathfrak{M}_B f > b_i$. Then we have

$$\sum_{B \in H_i} |B| = \sum_{B \in H_i} \int_B f |B| \Big/ \int_B f = \sum_{B \in H_i} \int_B f / \mathfrak{M}_B f < (1/b_i) \int_E |f|,$$

which tends to zero with i.

In case $\sup \mathfrak{R}(f) = b < \infty$, $b \bar{\in} \mathfrak{R}(f)$, we have $\lim_{i \to \infty} b_i = b$. Let $H_i \subset F$ stand for the family of sets B defined by the condition $\mathfrak{M}_B f > b_i$, and let $\epsilon_i = b - b_i$ $(i=1, 2, 3, \cdots)$. For any particular i and $B \in H_i$ let

$$B_1 = B \mathop{E}_{x} [b - \epsilon_i < f(x)], \qquad B_2 = B \mathop{E}_{x} [f(x) < b - 2\epsilon_i],$$

$$B_3 = B \mathop{E}_{x} [b - 2\epsilon_i \leqq f(x) \leqq b - \epsilon_i],$$

so that we have

$$\begin{aligned} b - \epsilon_i < \mathfrak{M}_B f &= \frac{|B_1|}{|B|} \mathfrak{M}_{B_1} f + \frac{|B_2|}{|B|} \mathfrak{M}_{B_2} f + \frac{|B_3|}{|B|} \mathfrak{M}_{B_3} f \\ &< \frac{|B_1|}{|B|} b + \frac{|B_2|}{|B|} (b - 2\epsilon_i) + \frac{|B_3|}{|B|} (b - \epsilon_i) \\ &= b - \epsilon_i (2|B_2|/|B| + |B_3|/|B|), \end{aligned}$$

whence, since ϵ_i is >0,

$$2|B_2|/|B|+|B_3|/|B|<1,$$
$$2|B_2|+|B_3|<|B|=|B_1|+|B_2|+|B_3|,$$
$$|B_2|<|B_1|,$$

and

$$|B|<2|B_1|+|B_3|\leqq 2(|B_1|+|B_3|).$$

Hence we obtain

$$\sum_{B\in H_i}|B|\leqq 2\left|\underset{x}{E}\,[b-2\epsilon_i\leqq f(x)]\right|.$$

But from the inequality $\epsilon_{i+1}\leqq\epsilon_i$ $(i=1, 2, 3, \cdots)$ we have

$$\underset{x}{E}\,[b-2\epsilon_{i+1}\leqq f(x)]\subset\underset{x}{E}\,[b-2\epsilon_i\leqq f(x)] \qquad (i=1, 2, 3, \cdots)$$

which implies

$$0\leqq\lim_{i\to\infty}\sup_{P\in\Gamma_\delta(E)}\sum_{B\in H_i}|B|\leqq 2\lim_{i\to\infty}\left|\underset{x}{E}\,[b-2\epsilon_i\leqq f(x)]\right|$$
$$=2\left|\prod_{i=1}^{\infty}\underset{x}{E}\,[b-2\epsilon_i\leqq f(x)]\right|=2\left|\underset{x}{E}\,[b\leqq f(x)]\right|=0.$$

The cases $\inf\mathfrak{R}(f)=-\infty$ and $\inf\mathfrak{R}(f)=a>-\infty$, $a\bar{\in}\mathfrak{R}(f)$ are entirely similar to those already considered, so that the lemma may now be regarded as established.

(5.5) DEFINITION. *For $f\in L^*(E)$, $\phi\in C(\mathfrak{R}(f))$, $0\leqq\phi(y)$ when $y\in\mathfrak{R}(f)$, and $\sum_{i=1}^{\infty}J_i$ a normal expression for $\mathfrak{R}(f)$, we call the sequence ϕ_i $(i=1, 2, 3, \cdots)$, where for each i*

$$\phi_i(y)=\begin{cases}\phi(y) & \textit{for}\quad y\in J_i=[a_i, b_i],\\ \inf\limits_{y\leqq t\leqq a_i}\phi(t) & \textit{for}\quad a_i>y\in\mathfrak{R}(f),\\ \inf\limits_{b_i\leqq t\leqq y}\phi(t) & \textit{for}\quad b_i<y\in\mathfrak{R}(f),\end{cases}$$

a normal approximating sequence for ϕ on $\mathfrak{R}(f)$.

We note that if $\mathfrak{R}(f)$ is a bounded closed interval, ϕ_i is identical with ϕ for each i. It is essential to observe also the following properties of a normal approximating sequence:

$$\phi_i\in UC(\mathfrak{R}(f)),$$
$$\phi(y)-\phi_i(y)=0 \quad\text{for}\quad y\in J_i,$$
$$0\leqq\phi_i(y)\leqq\phi(y)\quad\text{for}\quad y\in\mathfrak{R}(f),\qquad (i=1, 2, 3, \cdots);$$
$$\lim_{i\to\infty}\phi_i(y)=\phi(y)\quad\text{for}\quad y\in\mathfrak{R}(f).$$

Since the set of points $E\ E_x[f(x)\bar{\epsilon}\mathfrak{R}(f)]$ is of measure zero, it follows that we have

$$(5.6)\qquad \left.\begin{array}{l} 0 \leqq \phi_i[f(x)] \leqq \phi[f(x)] \quad (i = 1, 2, 3, \cdots) \\ \lim_{i\to\infty} \phi_i[(fx)] = \phi[f(x)] \end{array}\right\} \quad \text{for almost all } x \in E.$$

(5.7) THEOREM. *The conditions* $f\in L^*(E)$ *and* $\phi\in C(\mathfrak{R}(f))$ *imply*

$$\left|\int_E \phi:f\right| \leqq \int_E |\phi:f| \leqq S_*(f, |\phi|, E) \leqq \infty.$$

Proof. Let ϕ_i $(i=1, 2, 3, \cdots)$ be a normal approximating sequence for $|\phi|$ on $\mathfrak{R}(f)$, and let H represent any essentially bounded measurable subset of E. Theorems (5.2) and (4.1) (viii) imply

$$\int_H \phi_i:f = S(f, \phi_i, H) = S_*(f, \phi_i, H) \leqq S_*(f, |\phi|, H) \leqq S_*(f, |\phi|, E) \leqq \infty \qquad (i = 1, 2, 3, \cdots);$$

and by aid of (5.6) and a well known theorem of Fatou we obtain

$$\int_H |\phi:f| \leqq \liminf_{i\to\infty} \int_H \phi_i:f,$$

whence

$$\int_H |\phi:f| \leqq S_*(f, |\phi|, E).$$

The desired conclusion then follows from the arbitrariness of H.

(5.8) THEOREM. *The conditions* $f\in L^*(E)$, $\phi\in UC(\mathfrak{R}(f))$, $0\leqq\phi(y)$ *for* $y\in\mathfrak{R}(f)$, *and* $S^*(f, \phi, E)<\infty$ *are sufficient for* (1.9).

Proof. If E is essentially bounded, the conclusion follows from Theorem (5.2) without use of the hypothesis $S^*(f, \phi, E)<\infty$. Otherwise, this hypothesis implies the existence of η satisfying the conditions

$$(5.9)\qquad 0 < \eta < \infty, \qquad \sup_{F\in\Gamma_\eta(E)} \sum_{B\in F} \phi[\mathfrak{M}_B f]\,|B| < \infty.$$

In accordance with Theorem (4.3) let $N_1 = N_1(\epsilon/3k) > 0$ insure

$$\sup_{F\in\Gamma_\eta(H)} \sum_{B\in F} \phi[\mathfrak{M}_B f]\,|B| < \epsilon/(3k) \qquad \text{for } H\subset E - [-N_1, N_1].$$

From Theorem (5.7) we infer $\int_E\phi:f<\infty$; thus there exists N with

$$N_1 \leqq N < \infty, \qquad \int_H \phi:f < \epsilon/3 \qquad \text{for } H\subset E - [-N, N].$$

Let

$$E_N = E \underset{x}{E}\, [-N-\eta \leqq x \leqq N+\eta],$$

and in accordance with Theorem (5.2) let $\delta=\delta(f, \phi, E_N, \epsilon/3)$, $0<\delta\leqq\eta$, correspond to $\epsilon/3$ for the functions f and ϕ on E_N. For $F\in\Gamma_\delta(E)$ let $F_1\subset F$ stand for the family of sets $B\in F$ defined by the condition $B\subset E_N$, and let $F_2=F-F_1$. Then we have, by aid of Theorem (4.1) (iii),

$$\begin{aligned}
&\sum_{B\in F}\left|\int_B \phi:f - \phi[\mathfrak{M}_{AB}f]\,|B|\right| \\
&\qquad \leqq \sum_{B\in F_1}\left|\int_B \phi:f - \phi[\mathfrak{M}_{AB}f]\,|B|\right| + \sum_{B\in F_2}\int_B \phi:f + \sum_{B\in F_2}\phi[\mathfrak{M}_{AB}f]\,|B| \\
&\qquad < \epsilon/3 + \epsilon/3 + k \sup_{F\in\Gamma_\delta(H)} \phi[\mathfrak{M}_B f]\,|B| < \epsilon,
\end{aligned}$$

where $H=\sigma(F_2)\subset E-[-N_1, N_1]$.

For the proof of the next theorem the following lemma will prove helpful.

(5.10) LEMMA. *Let $f\in L^*(E)$, $\phi\in C(\mathfrak{R}(f))$, $0\leqq\phi(y)$ for $y\in\mathfrak{R}(f)$, $S^*(f, \phi, E)<\infty$, and ϕ_i $(i=1, 2, 3, \cdots)$ be a normal approximating sequence for ϕ on $\mathfrak{R}(f)$. Then there exists η with $0<\eta<\infty$ such that we have*

$$\lim_{i\to\infty}\sup_{F\in\Gamma_\eta(E)}\sum_{B\in F}(\phi-\phi_i)[\mathfrak{M}_{AB}f]\,|B| = 0.$$

Proof. If $\mathfrak{R}(f)$ is a bounded closed interval, $\phi(y)-\phi_i(y)=0$ for $y\in\mathfrak{R}(f)$ $(i=1, 2, 3, \cdots)$ and the conclusion is obvious with any η without use of the hypothesis $S^*(f, \phi, E)<\infty$. Otherwise, this hypothesis implies the existence of η satisfying conditions (5.9). In accordance with Theorem (4.3) let $N=N(\epsilon/2k)>\eta>0$ be such that

$$\sup_{F\in\Gamma_\eta(H)}\sum_{B\in F}\phi[\mathfrak{M}_B f]\,|B| < \epsilon/(2k) \qquad \text{for } H\subset E-[-N, N],$$

and in accordance with Theorem (4.4) let $\gamma=\gamma(f, \phi, \eta, \epsilon/(2k))>0$ be such that

$$H\subset E, \qquad |H|<\gamma \quad \text{imply} \quad \sup_{F\in\Gamma_\eta(H)}\sum_{B\in F}\phi[\mathfrak{M}_B f]\,|B| < \epsilon/(2k).$$

Let $E_N=[-N-\eta, N+\eta]E$ and in accordance with Lemma (5.4) let m be such that

$$\sup_{F\in\Gamma_\eta(E_N)}\sum_{B\in G_i}|B| < \gamma \qquad \text{for } i>m,$$

where $G_i\subset F$ has the meaning ascribed to it in that lemma. For $F\in\Gamma_\eta(E)$ let $F_1\subset F$ be defined by the condition $B\in F_1$ if and only if $[-N, N]B\neq 0$, and let $F_2=F-F_1$. For each i we then have, by aid of Theorem (4.1) (iii),

$$\sup_{F\in\Gamma_\eta(E)} \sum_{B\in F} (\phi - \phi_i)[\mathfrak{M}_{AB}f]\,|B| \leqq k \sup_{F\in\Gamma_\eta(E)} \sum_{B\in F} (\phi - \phi_i)[\mathfrak{M}_B f]\,|B|$$

$$\leqq k \sup_{F_1} \sum_{B\in F_1} (\phi - \phi_i)[\mathfrak{M}_B f]\,|B| + k \sup_{F_2} \sum_{B\in F_2} (\phi - \phi_i)[\mathfrak{M}_B f]\,|B|,$$

where the second term does not exceed

$$k \sup_{F_2} \sum_{B\in F_2} \phi[\mathfrak{M}_B f]\,|B| < k\epsilon/(2k) = \epsilon/2$$

and for $i > m$ the first term is less than or equal to

$$k \sup_{F\in\Gamma_\eta(E_N)} \sum_{B\in F} (\phi - \phi_i)[\mathfrak{M}_B f]\,|B| = k \sup_{F\in\Gamma_\eta(E_N)} \sum_{B\in G_i} (\phi - \phi_i)[\mathfrak{M}_B f]\,|B|$$

$$\leqq k \sup_{F\in\Gamma_\eta(E_N)} \sum_{B\in G_i} \phi[\mathfrak{M}_B f]\,|B| < k\epsilon/(2k) = \epsilon/2.$$

COROLLARY. *The hypotheses of this lemma imply*

$$\lim_{i\to\infty} S^*(f, \phi - \phi_i, E, k) = 0.$$

(5.11) THEOREM. *If $f\in L^*(E)$, $\phi\in C(\mathfrak{R}(f))$, and $0\leqq\phi(y)$ for $y\in\mathfrak{R}(f)$, the condition $S^*(f, \phi, E) < \infty$ is sufficient (as well as obviously necessary) for* (1.9).

Proof. If $\mathfrak{R}(f)$ is a bounded closed interval, the conclusion follows at once from Theorem (5.8). Otherwise, the hypothesis $S^*(f, \phi, E) < \infty$, by Theorem (5.7), implies the existence of $\int_E \phi : f$; and it also implies the existence of η satisfying conditions (5.9). Let ϕ_i ($i = 1, 2, 3, \cdots$) be a normal approximating sequence for ϕ on $\mathfrak{R}(f)$. Then for $0 < \delta \leqq \eta$, $F\in\Gamma_\delta(E)$, and each i we may write

$$\sum_{B\in F} \left| \int_B \phi : f - \phi[\mathfrak{M}_{AB}f]\,|B| \right|$$

$$\text{(a)} \qquad \leqq \int_E |\phi : f - \phi_i : f| + \sum_{B\in F} \left| \int_B \phi_i : f - \phi_i[\mathfrak{M}_{AB}f]\,|B| \right|$$

$$+ \sum_{B\in F} |\phi_i[\mathfrak{M}_{AB}f] - \phi[\mathfrak{M}_{AB}f]|\,|B|.$$

From (5.6) and the inequality $\int_E \phi : f < \infty$ we infer, by Lebesgue's convergence theorem, the existence of a number N_1 with

$$\int_E |\phi : f - \phi_i : f| < \epsilon/3 \qquad \text{for } i > N_1.$$

Lemma (5.10) implies the existence of a number $N \geqq N_1$ with

$$\sup_{F\in\Gamma_\delta(E)} \sum_{B\in F} |\phi_i[\mathfrak{M}_{AB}f] - \phi[\mathfrak{M}_{AB}f]|\,|B|$$

$$= \sup_{F\in\Gamma_\delta(E)} \sum_{B\in F} (\phi - \phi_i)[\mathfrak{M}_{AB}f]\,|B|$$

$$\leqq \sup_{F\in\Gamma_\eta(E)} \sum_{B\in F} (\phi - \phi_i)[\mathfrak{M}_{AB}f]\,|\,B\,| < \epsilon/3 \quad \text{for } i > N.$$

Having fixed $i>N$, in accordance with Theorem (5.8) we let $\delta=\delta\ (f, \phi_i, \epsilon/3)$, $0<\delta\leqq\eta$, be such that the second term on the right in (a) is $<\epsilon/3$. This completes the proof.

(5.12) THEOREM. *If $f\in L^*(E)$ and $\phi\in C(\mathfrak{R}(f))$, the condition*

$$-\infty < S_*(f, \phi, E) \leqq S^*(f, \phi, E) < \infty$$

is sufficient (as well as obviously necessary) for (1.9).

Proof. By the corollary to Theorem (4.2), the hypotheses of the present theorem imply $S^*(f, |\phi|, E)<\infty$. For $y\in\mathfrak{R}(f)$ let

$$\phi_1(y) = [\,|\,\phi(y)\,| + \phi(y)]/2, \qquad \phi_2(y) = [\,|\,\phi(y)\,| - \phi(y)]/2,$$

whence

$$\phi(y) = \phi_1(y) - \phi_2(y),$$

and for $i=1, 2$ we have

$$\phi_i \in C(\mathfrak{R}(f)),\ 0 \leqq \phi_i(y) \leqq |\,\phi(y)\,| \text{ for } y \in \mathfrak{R}(f),\ S^*(f, \phi_i, E) \leqq S^*(f, |\phi|, E) < \infty.$$

In accordance with Theorem (5.11), therefore, let $\delta_i=\delta_i\ (f, \phi_i, \epsilon/2)>0$ correspond to $\epsilon/2$ for the functions f and ϕ_i ($i=1, 2$). Then we have, for $\delta=\min\ [\delta_1, \delta_2]$ and $F\in\Gamma_\delta(E)$,

$$\begin{aligned}
&\sum_{B\in F}\left|\int_B \phi:f - \phi[\mathfrak{M}_{AB}f]\,|\,B\,|\right| \\
&\quad= \sum_{B\in F}\left|\int_B \phi_1:f - \int_B \phi_2:f - \phi_1[\mathfrak{M}_{AB}f]\,|\,B\,| + \phi_2[\mathfrak{M}_{AB}f]\,|\,B\,|\right| \\
&\quad\leqq \sum_{B\in F}\left|\int_B \phi_1:f - \phi_1[\mathfrak{M}_{AB}f]\,|\,B\,|\right| + \sum_{B\in F}\left|\int_B \phi_2:f - \phi_2[\mathfrak{M}_{AB}f]\,|\,B\,|\right| < \epsilon.
\end{aligned}$$

The following corollary is now evident.

COROLLARY. *For $f\in L^*(E)$ and $\phi\in C(\mathfrak{R}(f))$, the conditions* (1.8), (1.9), (1.10), *and* (1.11) *are equivalent.*

(5.13) THEOREM. *If $f\in L^*(E)$ and $\phi\in C(\mathfrak{R}(f))$, the conditions $|\phi(y)|\leqq\psi(y)$ for $y\in\mathfrak{R}(f)$, ψ convex on $\mathfrak{R}(f)$, and $\int_B\psi:f<\infty$ are sufficient for*

$$\text{(a)}\qquad \sup_{F\in\Gamma_\eta(E)}\left|\sum_{B\in F}\phi[\mathfrak{M}_B f]\,|\,B\,|\right| < \infty, \qquad \eta = \operatorname{diam} E,$$

and therefore for (1.9).

Proof. Using Jensen's inequality ([6]) we obtain, for $F\in\Gamma_\epsilon(E)$,

$$\left|\sum_{B\in F}\phi[\mathfrak{M}_Bf]\,|B|\right| \leqq \sum_{B\in F}|\phi[\mathfrak{M}_Bf]|\,|B| \leqq \sum_{B\in F}\psi[\mathfrak{M}_Bf]\,|B|$$

$$\leqq \sum_{B\in F}|B|\,\mathfrak{M}_B\psi:f = \int_B \psi:f < \infty,$$

from which follows $-\infty < S_*(f,\phi,E) \leqq S^*(f,\phi,E) < \infty$, the condition hypothesized in Theorem (5.12).

It will be shown in the next three sections that the conditions specified in this theorem are also necessary for (a), but are not necessary for (1.9).

Corollary. *If $f\in L^*(E)$, $\phi\in C(\mathfrak{R}(f))$ and $0\leqq\phi(y)$ for $y\in\mathfrak{R}(f)$, the condition ϕ convex on $\mathfrak{R}(f)$ implies*

$$\int_E \phi:f = S(f,\phi,E,k) \leqq \infty.$$

Proof. The function $\phi:f$ is measurable and non-negative on E. In case $\int_E\phi:f=\infty$, Theorem (5.7) implies $S(f,\phi,E,k)=S(f,\phi,E)=\infty$. In case $\int_E\phi:f<\infty$, the equality follows from Theorem (5.13).

6. **Some lemmas.** *Throughout this section we assume* that E is a measurable linear set with

$$|E| = \mu, \qquad 0<\mu<\infty;$$

$f\in L^*(E)$; ϕ is a non-negative function whose domain includes $\mathfrak{R}(f)$; and

$$\sup_{F\in\Gamma_\infty(E)}\sum_{B\in F}\phi[\mathfrak{M}_Bf]\,|B| = M < \infty.$$

We begin by developing a few properties of equimeasurable functions which are needed for our special purposes and some of which may not be immediately accessible in the literature.

(6.1) Definition. *Of two functions u and v we say that u on A is equimeasurable with v on B if and only if A and B are sets included, respectively, in the domains of u and v and the sets*

$$A\,\underset{t}{E}\,[u(t)\leqq y], \qquad B\,\underset{s}{E}\,[v(s)\leqq y]$$

are measurable and of equal measure for $-\infty<y<\infty$.

We now set

([6]) See, for example, Hardy, Littlewood, and Pólya, *Inequalities*, Cambridge, 1934, pp. 150–152. The reader should bear in mind, here and later on, that the condition ψ convex on an interval implies ψ continuous on the interior of that interval.

$$\alpha(y) = \left| \underset{t}{E}\, [f(t) \leqq y] \right| \qquad \text{for } -\infty < y < \infty,$$

and on the open interval $(0, \mu)$ define the function g by the relation

$$g(s) = \inf \underset{y}{E}\, [s \leqq \alpha(y)], \qquad 0 < s < \mu.$$

(6.2) LEMMA. *On $(0, \mu)$ the function g is non-decreasing and equimeasurable with f on E.*

Proof. Obviously g is non-decreasing. Let $-\infty < y_0 < \infty$, let $a = \alpha(y_0)$, and set

$$b = \left| \underset{s}{E}\, [g(s) \leqq y_0] \right|.$$

The following properties may then be readily deduced seriatim:

(a) $\qquad s \leqq \alpha(g(s)) \qquad$ for $0 < s < \mu$;

(b) $\qquad g(s) \leqq y_0$ implies $s \leqq \alpha(g(s)) \leqq \alpha(y_0) = a$;

(c) $\qquad g(s) > y_0$ implies $s > a$.

Since g is non-decreasing on $(0, \mu)$ we have

$$\sup \left(\{0\} + \underset{s}{E}\, [g(s) \leqq y_0] \right) = b = \inf \left(\{\mu\} + \underset{s}{E}\, [g(s) > y_0] \right);$$

this combines with (b) to yield $b \leqq a$ and with (c) to provide $b \geqq a$, whence $a = b$ and the lemma is proved.

(6.3) LEMMA. *If D is an open interval and D' a measurable set with $D \subset D' \subset (0, \mu)$, and if B' is a subset of E such that f on B' is equimeasurable with g on D', then there exists a set $B \subset B'$ such that f on B is equimeasurable with g on D and f on $B' - B$ is equimeasurable with g on $D' - D$.*

Proof. Let $(a, b) = D$, and let

$$a' = \inf_{a<t<b} g(t), \qquad b' = \sup_{a<t<b} g(t),$$

$$B_1 = B' \underset{t}{E}\, [f(t) = a'], \qquad D_1 = D \underset{s}{E}\, [g(s) = a'],$$

$$B_2 = B' \underset{t}{E}\, [a' < f(t) < b'], \qquad D_2 = D \underset{s}{E}\, [a' < g(s) < b'],$$

$$B_3 = B' \underset{t}{E}\, [a' < f(t) = b'], \qquad D_3 = D \underset{s}{E}\, [a' < g(s) = b'].$$

A simple check shows that

$$D_2 = D' \underset{s}{E}\, [a' < g(s) < b']$$

and that f on B_2 is equimeasurable with g on D_2. Clearly we have $|D_1| \leqq |B_1|$ and $|D_3| \leqq |B_3|$. By considering the sets

$$(-x, x)B_1 \quad \text{and} \quad (-x, x)B_3 \qquad \text{for } 0 < x,$$

we find measurable sets β_1 and β_3 with

$$\beta_1 \subset B_1, \qquad \beta_3 \subset B_3, \qquad |\beta_1| = |D_1|, \qquad |\beta_3| = |D_3|.$$

Thus f on β_1 is equimeasurable with g on D_1 and f on β_3 is equimeasurable with g on D_3. Accordingly, on account of the disjointness of the sets involved, we may take $B = \beta_1 + B_2 + \beta_3$, note that $D = D_1 + D_2 + D_3$, and conclude that f on B is equimeasurable with g on D. It then follows at once that f on $B' - B$ is equimeasurable with g on $D' - D$.

(6.4) Lemma. *If u on A is equimeasurable with v on B, then*

$$\int_A u = \int_B v \quad \text{and} \quad |A| = |B|.$$

This well known result is easily seen.

(6.5) Lemma. *If $\mathfrak{G}$ is a finite disjointed family of open subintervals of* $(0, \mu)$, *we have*

$$\sum_{D \in \mathfrak{G}} \phi[\mathfrak{M}_D g]\,|D| \leqq M.$$

Proof. Let the number of intervals in $\mathfrak{G}$ be n and let

$$\mathfrak{G} = \{D_1\} + \{D_2\} + \cdots + \{D_n\}.$$

By repeated use of Lemma (6.3) we infer the existence of a disjointed family F of n subsets of E such that

$$F = \{B_1\} + \{B_2\} + \cdots + \{B_n\},$$

with f on B_i equimeasurable with g on D_i $(i = 1, 2, \cdots, n)$. From Lemma (6.4) we have

$$\mathfrak{M}_{B_i} f = \mathfrak{M}_{D_i} g, \qquad |B_i| = |D_i| \qquad (i = 1, 2, \cdots, n),$$

whence

$$\begin{aligned}\sum_{D \in \mathfrak{G}} \phi[\mathfrak{M}_D g]\,|D| &= \sum_{i=1}^{n} \phi[\mathfrak{M}_{D_i} g]\,|D_i| = \sum_{i=1}^{n} \phi[\mathfrak{M}_{B_i} f]\,|B_i| \\ &= \sum_{B \in F} \phi[\mathfrak{M}_B f]\,|B| \leqq M.\end{aligned}$$

(6.6) Lemma. *We have* $\mathfrak{R}(f) = \mathfrak{R}(g)$.

This result is almost immediate.

(6.7) LEMMA. *If (a, b) is an open subinterval of $(0, \mu)$ and if ψ is a function linear on an interval containing the image of (a, b) under g with*

$$\psi[\mathfrak{M}_{(a,b)}g] = \phi[\mathfrak{M}_{(a,b)}g],$$

then we have

$$\int_a^b \psi:g = \phi[\mathfrak{M}_{(a,b)}g](b-a).$$

Proof. Let J be an interval which contains the image of (a, b) under g and on which ψ is linear. We have

$$\psi(y) = \omega y - \omega\gamma + \phi(\gamma) \qquad \text{for } y \in J,$$

where ω is a suitable number and $\gamma = \mathfrak{M}_{(a,b)}g$, whence

$$\int_a^b \psi:g = \int_a^b (\omega g(s) - \omega\gamma + \phi(\gamma))ds = \omega\int_a^b g - \omega\gamma(b-a) + \phi(\gamma)(b-a)$$
$$= \phi(\gamma)(b-a).$$

(6.8) DEFINITION. *A family K is irreducible if and only if*

$$\sigma(K - \{C\}) \neq \sigma(K) \qquad \textit{for } C \in K.$$

(6.9) DEFINITION. *A finite irreducible family of open intervals is called a chain if and only if $\sigma(K)$ is an open interval.*

Concerning chains the following remarks may be in order. No interval in a chain is the vacuous set, but the vacuous set itself constitutes a chain. The intervals in a chain may be ordered according to non-decreasing left-hand [right-hand] endpoints; these two orders are the same. Taking alternate intervals in this ordering one readily sees that a chain is expressible as the sum of two disjoint families each of which is itself disjointed. If K is a chain, $K_1 \subset K$, and $\sigma(K_1)$ is an interval, then K_1 is a chain. If the interval farthest to the left [right] in a chain is deleted, the remaining family is a chain. If K is a chain and C is an open interval which overlaps $\sigma(K)$ but contains no element of K, $K + \{C\}$ is a chain.

(6.10) **Abbreviations.** Let η be a positive number fixed throughout the rest of this section. We set

$$W(D) = (\eta/2\mu)|D| + \phi[\mathfrak{M}_D g]|D|,$$
$$\mathfrak{R}^* = \{\mathfrak{M}_{(0,\mu)}g\} + \text{the interior of } \mathfrak{R}(g).$$

For n a positive integer we employ $\langle n\rangle$ as an abbreviation for the following statement:

If H is a set with n elements and if

$$\mathfrak{M}_{(0,\mu)}g \in H \subset \mathfrak{R}^*,$$

then there exists a non-negative convex function ψ on $-\infty<y<\infty$ which dominates ϕ on H, a chain K with $\sigma(K)=(0, \mu)$, and for each $D\in K$ a non-negative function F_D on $(0, \mu)$, all of which are so related that we have

$$\int_0^\mu F_D < W(D) \qquad \text{for } D \in K,$$

and $0<s<\mu$ implies the existence of a $D\in K$ for which these conditions are satisfied:

$$s \in D, \qquad \psi[g(s)] = F_D(s).$$

(6.11) LEMMA. *$\langle 1\rangle$ is valid.*

Proof. Let $C=(0, \mu)$, $H=\{\mathfrak{M}_C g\}$, and ψ be the non-negative convex function on $(-\infty, \infty)$ defined by

$$\psi(y) = \phi[\mathfrak{M}_C g] \qquad \text{for } -\infty < y < \infty.$$

Let $K=\{C\}$ and define F_C on $(0, \mu)$ by

$$F_C(s) = \phi[\mathfrak{M}_C g] \qquad \text{for } 0 < s < \mu.$$

We then have

$$\int_0^\mu F_C = \mu\phi[\mathfrak{M}_C g] = \phi[\mathfrak{M}_C g]\,|\,C\,| < W(C),$$

and $0<s<\mu$ implies

$$F_C(s) = \phi[\mathfrak{M}_C g] = \psi[g(s)].$$

(6.12) LEMMA. *If ϕ is continuous on $\mathfrak{R}(f)$ and n is a positive integer, the validity of $\langle n\rangle$ implies that of $\langle n+1\rangle$.*

Proof. Let H be a set with $n+1$ elements and assume

$$\mathfrak{M}_{(0,\mu)} g \in H \subset \mathfrak{R}^*.$$

Either there is a number in H which is $>\mathfrak{M}_{(0,\mu)}g$ or there is a number in H which is $<\mathfrak{M}_{(0,\mu)}g$. For definiteness we assume henceforth that the former is the case and we designate by ν the largest such number. It should be clear from the nature of the ensuing argument that analogous reasoning can be brought to bear on the alternative case.

Let

$$\overline{H} = H - \{\nu\},$$

so that $\overline{H}$ is a set of n elements and

$$\mathfrak{M}_{(0,\mu)} g \in \overline{H} \subset \mathfrak{R}^*.$$

In view of the validity of $\langle n\rangle$, let $\bar{\psi}$ be a non-negative convex function on $(-\infty, \infty)$ which dominates ϕ on $\overline{H}$; let $\overline{K}$ be a chain with $\sigma(\overline{K})=(0, \mu)$;

and for each $D \in \overline{K}$ let $\overline{F}_D$ be a non-negative function on $(0, \mu)$, all of which are so related that we have

(a) $$\int_0^\mu F_D < W(D) \qquad \text{for } D \in \overline{K},$$

and

(b) $0 < s < \mu$ implies the existence of a $D \in \overline{K}$ with

$$s \in D, \quad \overline{\psi}[g(s)] = \overline{F}_D(s).$$

In the case $\phi(\nu) \leqq \overline{\psi}(\nu)$ it is apparent that $\langle n+1 \rangle$ holds. From now on we assume

(c) $$\overline{\psi}(\nu) < \phi(\nu).$$

Recalling that $\mathfrak{M}_{(0,\mu)}g$ is $<\nu$ we let ξ satisfy the conditions

(d) $$0 < \xi < \mu, \quad \mathfrak{M}_{(\xi,\mu)}g = \nu.$$

Let C_1 be the interval farthest to the right in $\overline{K}$. By moving only the right endpoint of C_1 a suitably small amount to the left we obtain an interval C_2 which not only enjoys the property that its right endpoint is interior to the interval

$$(\xi, \mu) - \sigma(\overline{K} - \{C_1\})$$

but also, in view of the continuity of ϕ, the property

$$\int_0^\mu \overline{F}_{C_1} < W(C_2).$$

Recalling (a) and defining

(e) $$\overline{F}_{C_2} = \overline{F}_{C_1}, \quad \overline{K}_1 = \overline{K} - \{C_1\} + \{C_2\},$$

we see at once

(f) $$\int_0^\mu \overline{F}_D < W(D) \qquad \text{for } D \in \overline{K}_1.$$

Our determination of C_2 insures that $\overline{K}_1$ is a chain in which the interval farthest to the right is C_2, whence

(g) $$s \in \sigma(\overline{K}_1)C_1 \quad \text{implies} \quad s \in C_2.$$

Making use of (b), the first relation in (e), and (g) we infer

(h) $s \in \sigma(\overline{K}_1)$ implies the existence of a $D \in \overline{K}_1$ with

$$s \in D, \quad \overline{\psi}[g(s)] = \overline{F}_D(s).$$

Setting

$$\overline{K}_2 = \overline{K}_1 - \overline{K}_1 \underset{D}{E} [D \subset (\xi, \mu)],$$

$$K = \overline{K}_2 + \{(\xi, \mu)\},$$

we note that $\overline{K}_2$ and K are chains and $\sigma(K) = (0, \mu)$. Since ν is interior to $\Re(g)$ we obtain from (d) the inequality $g(\xi) < \nu$. Clearly there exists a function ψ on $(-\infty, \infty)$ with

$$\text{(i)} \qquad \begin{array}{ll} \psi(y) = \bar{\psi}(y) & \text{for } y \leqq g(\xi), \\ \psi \text{ linear on } \underset{y}{E} [g(\xi) \leqq y], \quad \psi(\nu) = \phi(\nu). & \end{array}$$

Let

$$\text{(j)} \qquad F_D = \overline{F}_D \qquad \text{for } D \in \overline{K}_2;$$

for $D = (\xi, \mu)$ let F_D be defined on $(0, \mu)$ by the relations

$$\text{(k)} \qquad \begin{array}{ll} F_D(s) = \psi[g(s)] & \text{for } \xi < s < \mu, \\ F_D(s) = 0 & \text{for } 0 < s \leqq \xi. \end{array}$$

By aid of (c) we find

$$\bar{\psi}(\nu) < \phi(\nu) = \psi(\nu);$$

from this relation and (i) it follows that ψ is non-negative and convex on $(-\infty, \infty)$ with

$$\bar{\psi}(y) \leqq \psi(y) \qquad \text{for } y \leqq \nu.$$

Thus ψ dominates ϕ not only on $\overline{H}$, since each number in $\overline{H}$ is $<\nu$, but on H as well.

From (k), (d), (i), Lemma (6.7), and Abbreviations (6.10) we see that for $D = (\xi, \mu)$ we have

$$\int_0^\mu F_D = \int_\xi^\mu F_D = \int_\xi^\mu \psi : g = \phi[\mathfrak{M}_{(\xi,\mu)} g](\mu - \xi) = \phi[\mathfrak{M}_D g] \, | D | < W(D).$$

For $D = (\xi, \mu)$ the validity of the relation

$$\text{(l)} \qquad \int_0^\mu F_D < W(D)$$

is thus evident; that it is valid for $D \in \overline{K}_2$ is clear from (f) and the relation $\overline{K}_2 \subset \overline{K}_1$. From the definition of K we then have relation (l) for $D \in K$.

For $0 < s \leqq \xi$ we have $s \in \sigma(\overline{K}_2) \subset \sigma(\overline{K}_1)$, and by (h) there exists $D \in \overline{K}_1$ with

$$s \in D, \qquad \bar{\psi}[g(s)] = \overline{F}_D(s);$$

for this particular D it is clear that we have

$$D \in \overline{K}_2 \subset K, \qquad g(s) \leqq g(\xi),$$

and hence, by aid of (i) and (j),

$$\text{(m)} \qquad s \in D, \quad \psi[g(s)] = F_D(s).$$

The proof that $0<s<\mu$ implies the existence of a $D \in K$ with property (m) is now readily completed by reference to (k).

The proof of the lemma is now complete.

(6.13) LEMMA. *If ϕ is continuous on $\Re(f)$, H is a finite subset of the interior of $\Re(f)$, and η is an arbitrary positive number, there exists a non-negative convex function ψ on $(-\infty, \infty)$ which dominates ψ on H and satisfies the condition*

$$\int_E \psi : f < \eta + 2M.$$

Proof. Let $H^* = H + \{\mathfrak{M}_E f\}$. From Lemmas (6.4) and (6.6) we have

$$\mathfrak{M}_{(0,\mu)} g = \mathfrak{M}_E f, \qquad \Re(g) = \Re(f);$$

hence, from the definition of $\Re^*$ in (6.10), we see that H^* is a finite set with

$$\mathfrak{M}_{(0,\mu)} g \in H^* \subset \Re^*.$$

Lemmas (6.11) and (6.12) insure the existence of a non-negative convex function ψ which dominates ϕ on H^*, of a chain K with $\sigma(K) = (0, \mu)$, and for each $D \in K$ a non-negative function F_D on $(0, \mu)$, all so related as to have the properties specified in $\langle n \rangle$ under (6.10):

$$\text{(a)} \qquad \int_0^\mu F_D < W(D) \qquad \text{for } D \in K,$$

(b) $\qquad 0 < s < \mu$ implies the existence of a $D \in K$ with

$$s \in D, \qquad \psi[g(s)] = F_D(s).$$

It was pointed out in the remarks following Definition (6.9) that we may write

$$K = \mathfrak{G}_1 \dotplus \mathfrak{G}_2,$$

where $\mathfrak{G}_1$ is a disjointed family of intervals in K, $\mathfrak{G}_2$ is likewise, and $\mathfrak{G}_1\mathfrak{G}_2$ is vacuous. Thus, using (6.10) and Lemma (6.5), we have

$$\begin{aligned}
\sum_{D\in K} W(D) &= \sum_{D\in\mathfrak{G}_1} W(D) + \sum_{D\in\mathfrak{G}_2} W(D) \\
&\leqq (\eta/2\mu) \sum_{D\in\mathfrak{G}_1} |D| + (\eta/2\mu) \sum_{D\in\mathfrak{G}_2} |D| + \sum_{D\in\mathfrak{G}_1} \phi[\mathfrak{M}_D g] \, |D| \\
&\quad + \sum_{D\in\mathfrak{G}_2} \phi[\mathfrak{M}_D g] \, |D| \\
&\leqq (\eta/2\mu)\mu + (\eta/2\mu)\mu + M + M = \eta + 2M.
\end{aligned}$$

The relation

$$\psi[g(s)] \leqq \sum_{D\in K} F_D(s) \qquad \text{for } 0 < s < \mu$$

is an immediate consequence of (b) and the non-negativity of F_D. These last two inequalities now combine with (a) to yield

$$\text{(c)} \qquad \int_0^\mu \psi:g \leqq \int_0^\mu \Big(\sum_{D\in K} F_D\Big) = \sum_{D\in K} \int_0^\mu F_D < \sum_{D\in K} W(D) \leqq \eta + 2M.$$

Since g on $(0, \mu)$ is equimeasurable with f on E and $E_y[\psi(y) \leqq a]$ is an interval for $-\infty < a < \infty$, it is readily shown that $\psi:g$ on $(0, \mu)$ is equimeasurable with $\psi:f$ on E. Accordingly, from Lemma (6.3) and (c) we have

$$\int_E \psi:f = \int_0^\mu \psi:g < \eta + 2M.$$

Since ψ dominates ϕ on H^*, it dominates ϕ on H. The proof is complete.

(6.14) LEMMA. *If J is an interval and for each positive integer n the function ψ_n is non-negative and convex on J with*

$$\limsup_{n\to\infty} \psi_n(t) < \infty \qquad \textit{for } t \in J,$$

then there exists a function ψ_0 non-negative and convex on J and integers N_i $(i=1, 2, 3, \cdots)$ with the properties

$$1 \leqq N_1 < N_2 < N_3 < \cdots, \quad \lim_{i\to\infty} \psi_{N_i}(t) = \psi_0(t) \qquad \textit{for } t \in J.$$

Proof. Let $J_m = [a_m, b_m]$ $(m=1, 2, 3, \cdots)$ be a sequence of closed intervals with

$$\text{(a)} \qquad J_m \subset J_{m+1} \quad (m = 1, 2, 3, \cdots), \qquad \sum_{m=1}^{\infty} J_m = J;$$

and let

$$C = \sum_{n=1}^{\infty} \{\psi_n\}.$$

For each pair of positive integers m and n the total variation of ψ_n on J_m is less than or equal to $\psi_n(a_m)+\psi_n(b_m)$, a number which is also a bound for ψ_n on J_m. Thus there exists a sequence of non-negative numbers M_m $(m=1, 2, 3, \cdots)$ such that for each positive integer m and each $\psi \in C$ the total variation of ψ on J_m, and the values of ψ on J_m, do not exceed M_m. From a theorem of Helly(7) we conclude that for each positive integer m the

(7) Helly, *Über lineare Funktionaloperationen*, Sitzungsberichte der Akademie der Wissenschaften in Wien vol. 121, IIa (1912) p. 283.

family C is conditionally compact in the topology of pointwise convergence on J_m; that is, each sequence of elements of C contains a subsequence which is pointwise convergent on J_m.

Using a familiar diagonal process of selection and the properties (a) of the intervals J_m, we see that in the topology of pointwise convergence on J the family C is also conditionally compact. Hence there exist integers N_i $(i=1, 2, 3, \cdots)$ with $1 \leqq N_1 < N_2 < N_3 < \cdots$ such that

$$\lim_{i\to\infty} \psi_{N_i}(t) \text{ exists (finite)} \qquad \text{for } t \in J.$$

Defining ψ_0 on J as this limit, it follows at once that ψ_0 is non-negative and convex on J.

7. **Concerning convex dominants for the function ϕ.** In this section E is to be understood as a measurable linear set with

$$|E| = \mu, \qquad 0 < \mu \leqq \infty.$$

(7.1) THEOREM. *If $f \in L^*(E)$, ϕ is a non-negative continuous function on $\mathfrak{R}(f)$, and*

$$\sup_{F\in\Gamma_\infty(E)} \sum_{B\in F} \phi[\mathfrak{M}_B f]\,|B| = M < \infty,$$

then there exists a non-negative convex function ψ which dominates ϕ on $\mathfrak{R}(f)$, with

$$\int_E \psi : f \leqq 2M.$$

Proof. If the interior of $\mathfrak{R}(f)$ is vacuous, $\mathfrak{R}(f)$ consists of but a single point; in this case ψ may be taken as the non-negative convex function $\psi(y)=\phi(y)$ for $y \in \mathfrak{R}(f)$ and the conclusion is apparent. Accordingly we assume from now on that the interior of $\mathfrak{R}(f)$ is not vacuous.

Let $m>0$ be such that the condition

$$|[-m, m]E| > 0$$

is satisfied. For each positive integer n let

$$E_n = [-mn, mn]E;$$

let

$$\{r_1\} + \{r_2\} + \{r_3\} + \cdots$$

be the set of all rational numbers; let f_n be the function defined on E_n by

$$f_n(t) = f(t) \qquad \text{for } t \in E_n;$$

let

$$H_n = [\text{interior of } \mathfrak{R}(f_n)][\{r_1\} + \{r_2\} + \cdots + \{r_n\}];$$

and let h_n be the characteristic function of E_n. Finally, let

$$\text{(a)} \qquad H = \sum_{n=1}^{\infty} H_n.$$

We see at once the relations

$$\text{(b)} \qquad \lim_{n\to\infty} h_n(t) = 1 \qquad \text{for } t \in E,$$

$$\text{(c)} \qquad \mathfrak{R}(f_n) \subset \mathfrak{R}(f_{n+1}), \quad H_n \subset H_{n+1} \qquad (n = 1, 2, 3, \cdots).$$

A more careful check shows

$$\text{(d)} \qquad \mathfrak{R}(f) = \sum_{n=1}^{\infty} \mathfrak{R}(f_n),$$

$$\text{(e)} \qquad \overline{H} \supset \mathfrak{R}(f),$$

where $\overline{H}$ represents the closure of H.

For each positive integer n we have

$$\sup_{F\in\Gamma_\infty(E_n)} \sum_{B\in F} \phi[\mathfrak{M}_B f_n] \mid B \mid \leqq \sup_{F\in\Gamma_\infty(E)} \sum_{B\in F} \phi[\mathfrak{M}_B f] \mid B \mid = M,$$

and H_n is a finite subset of $\mathfrak{R}(f_n)$; by Lemma (6.13) there exists a non-negative convex function ψ_n on $(-\infty, \infty)$ which dominates ϕ on H_n and satisfies the inequality

$$\int_{E_n} \psi_n : f_n \leqq 2M + 1/n.$$

Let $y_0 \in \mathfrak{R}(f)$. Recalling (c) and (d) we let n_0 be such that $y_0 \in \mathfrak{R}(f_n)$ for $n \geqq n_0$. Let B_0 be a subset of E_{n_0} with $\mathfrak{M}_{B_0} f_{n_0} = y_0$. From the definition of f_n and the relations

$$E_n \subset E_{n+1} \qquad (n = 1, 2, 3, \cdots)$$

follows

$$y_0 = \mathfrak{M}_{B_0} f = \mathfrak{M}_{B_0} f_n \qquad \text{for } n \geqq n_0,$$

and, by aid of Jensen's inequality,

$$\psi_n(y_0) = \psi_n[\mathfrak{M}_{B_0} f_n] \leqq \mathfrak{M}_{B_0} \psi_n : f_n \leqq \frac{2M + 1/n}{\mid B_0 \mid} \qquad \text{for } n \geqq n_0,$$

whence

$$\limsup_{n\to\infty} \psi_n(y_0) \leqq 2M / \mid B_0 \mid < \infty.$$

By Lemma (6.14) we now infer the existence of a non-negative convex function ψ on $\mathfrak{R}(f)$ and a sequence of integers N_i $(i = 1, 2, 3, \cdots)$ with the properties

$$1 \leqq N_1 < N_2 < N_3 < \cdots, \quad \lim_{i\to\infty} \psi_{N_i}(y) = \psi(y) \qquad \text{for } y \in \mathfrak{R}(f).$$

Since $f(x)\in\mathfrak{R}(f)$ for almost all $x\in E$, it may be inferred by aid of (b) that

$$\lim_{i\to\infty} h_{N_i}(x)\psi_{N_i}[f(x)] = \lim_{i\to\infty} \psi_{N_i}[f(x)] = \psi[f(x)]$$

for almost all $x\in E$. Using a well known theorem of Fatou we then obtain

$$\begin{aligned}\int_E \psi:f &\leqq \liminf_{i\to\infty} \int_E h_i(x)\psi_{N_i}[f(x)]dx = \liminf_{i\to\infty} \int_{E_{N_i}} \psi_{N_i}:f \\ &= \liminf_{i\to\infty} \int_{E_{N_i}} \psi_{N_i}:f_{N_i} = \lim_{i\to\infty} \int_{E_{N_i}} \psi_{N_i}:f_{N_i} \\ &\leqq \lim_{i\to\infty} (2M + 1/N_i) = 2M.\end{aligned}$$

From the second part of (c) and (a) it follows that ψ dominates ϕ on H. From (e), the fact that the convex function ψ is upper semi-continuous on $\mathfrak{R}(f)$, and the continuity of ϕ, we now conclude that ψ dominates ϕ on $\mathfrak{R}(f)$.

Remark. In Theorem (7.1) the hypothesis that ϕ is continuous can be dispensed with. To show this in detail appears to be a tedious process, but it may be indicated roughly as follows. Employing properties of equimeasurable functions as given in §6, Theorem (7.1) may be used to establish this result: *If $f\in L^*(E)$, ϕ is a non-negative function on $\mathfrak{R}(f)$,*

$$\sup_{F\in\Gamma_\infty(E)} \sum_{B\in F} \phi[\mathfrak{M}_B f]\,|B| = M < \infty,$$

H is a finite subset of $\mathfrak{R}(f)$, and η is an arbitrary positive number, there exists a non-negative convex function ψ on $(-\infty, \infty)$ which dominates ϕ on H and satisfies the condition

$$\int_E \psi:f \leqq \eta + 2M.$$

Using this fact and suitably redefining H_n, as it was introduced in the proof of Theorem (7.1), we may reach the desired conclusion. It is then easy to remove from Theorem (7.3) below the hypothesis that ϕ is continuous.

(7.2) Theorem. *If $f\in L^*(E)$, ϕ is a non-negative continuous function on $\mathfrak{R}(f)$, and*

$$\sup_{F\in\Gamma_\infty(E)} \sum_{B\in F} \phi[\mathfrak{M}_B f]\,|B| = M < \infty,$$

there exists a non-negative continuous convex function ψ which dominates ϕ on $\mathfrak{R}(f)$, with

$$\int_E \psi:f \leqq 2M.$$

Proof. By Theorem (7.1) there exists a non-negative convex function ψ_0

which dominates ϕ on $\Re(f)$, with $\int_E\psi_0:f\leqq 2M$. If the interior of $\Re(f)$ is vacuous, ψ_0 may be taken as ψ. Otherwise let ψ be the continuous function on $\Re(f)$ which is identical with ψ_0 on the interior of $\Re(f)$. Since ϕ is continuous and ψ_0 is convex, we have

$$\phi(y) \leqq \psi(y) \leqq \psi_0(y) \qquad \text{for } y \in \Re(f)$$

and the conclusion is immediate.

(7.3) THEOREM. *If* $f\in L^*(E)$, $\phi\in C(\Re(f))$, *and*

$$\sup_{F\in\Gamma_\infty(E)}\left|\sum_{B\in F}\phi[\mathfrak{M}_Bf]\,|B|\right| < \infty,$$

there exists a non-negative continuous convex function ψ which dominates $|\phi|$ on $\Re(f)$, with

$$\int_E \psi:f < \infty.$$

Proof. Defining the function Φ on $\Re(f)$ by

$$\Phi(y) = |\phi(y)| \qquad \text{for } y \in \Re(f),$$

we infer from Theorem (3.12)

$$\sup_{F\in\Gamma_\infty(E)}\sum_{B\in F}\Phi[\mathfrak{M}_Bf]\,|B| < \infty.$$

The conclusion then follows from Theorem (7.2).

It is almost superfluous to remark that in Theorems (7.1)–(7.3), if the set E is essentially bounded, we have $L^*(E)=L(E)$ and $\Gamma_\infty(E)$ can be replaced by $\Gamma_\eta(E)$ where η represents the essential diameter of E.

8. **Certain counterexamples.** In order that the question at issue should be completely covered, the theorems of §7 above must be supplemented by an example to show that for $f\in L^*(E)$ and $\phi\in C(\Re(f))$ the condition that $|\phi|$ be dominated by a convex function as specified in Theorem (5.13) is not necessary for (1.9). In this section we shall give two examples and conclude with a few supplementary remarks.

EXAMPLE A. The following example exhibits a function $f\in L(I)$ and a non-negative function $\phi\in C(\Re(f))$ for which relation (1.9) holds but ϕ is dominated on $\Re(f)$ by no convex function ψ with $\int_I\psi:f<\infty$.

Let ϕ_1 be defined on $(-\infty,\infty)$ thus: for $-\infty<y\leqq 0$, $\phi_1(y)=0$; for $n\geqq 1$ and odd, $\phi_1(2n)=2^{(2n)^2}$, $\phi_1(2n+2)=2^{(2n+3)^2}$; for $n\geqq 2$ and even, $\phi_1(2n)=2^{(2n+1)^2}$, $\phi_1(2n+2)=2^{(2n+2)^2}$; on each closed interval $[2n, 2n+2]$ $(n=0, 1, 2, \cdots)$, ϕ_1 is linear. Let ϕ_2 be defined in the same manner except for interchange of the specifications when n is odd and when n is even. It may be verified at once that the slope of ϕ_i $(i=1, 2)$ on the interval $[2n, 2n+2]$ is an increasing

function of n; thus ϕ_i is convex on $(-\infty, \infty)$. For each $y\in(-\infty, \infty)$ let $\phi(y)=\min[\phi_1(y), \phi_2(y)]$, so that ϕ is continuous on $(-\infty, \infty)$.

Let I_n $(n=1, 2, 3, \cdots)$ be a sequence of disjoint intervals with $|I_n|=1/2^{(2n)^2+n}$, $I_n\subset[2/3, 1]$ for n odd, and $I_n\subset[0, 1/3]$ for n even. Let the functions f_1, f_2, and f be defined on I by

$$f_1(x)=\begin{cases}2n \text{ for } x\in I_n \text{ and } n \text{ odd,}\\ 0 \text{ otherwise,}\end{cases}\qquad f_2(x)=\begin{cases}2n \text{ for } x\in I_n \text{ and } n \text{ even,}\\ 0 \text{ otherwise,}\end{cases}$$

$$f(x)=f_1(x)+f_2(x)\qquad \text{for } x\in I.$$

One verifies at once $f_1\in L(I)$, $f_2\in L(I)$, and hence $f\in L(I)$.

Under the condition diam $(B)<\delta\leqq 1/3$, which we now assume to be in force, a set B which contains a point of either interval $[0, 1/3]$, $[2/3, 1]$ contains no point of the other. From the definitions of f_1 and f_2 it is therefore clear that

$$\int_{AB} f_1>0 \quad\text{implies}\quad \int_{AB} f_2=0$$

and vice versa. Thus, for $F\in\Gamma_\delta(I)$, we have

$$\begin{aligned}\sum_{B\in F}\phi[\mathfrak{M}_{AB}f]\,|B| &= \sum_{B\in F}\phi[\mathfrak{M}_{AB}f_1+\mathfrak{M}_{AB}f_2]\,|B|\\ &= \sum_{B\in F}\phi[\mathfrak{M}_{AB}f]\,|B|+\sum_{B\in F}\phi[\mathfrak{M}_{AB}f]\,|B|,\end{aligned}$$

as well as

$$\int_B\phi:f=\int_B\phi:(f_1+f_2)=\int_B\phi:f_1+\int_B\phi:f_2\qquad \text{for } B\in F.$$

We may then infer from Theorem (5.13)

$$\begin{aligned}&\lim_{\delta\to 0}\ \sup_{F\in\Gamma_\delta(I)}\ \sum_{B\in F}\big|\,\mathfrak{M}_B\phi:f-\phi[\mathfrak{M}_{AB}f]\,\big|\,|B|\\ &\quad=\lim_{\delta\to 0}\ \sup_{F\in\Gamma_\delta(I)}\ \sum_{B\in F}\big|\,\mathfrak{M}_B\phi:f_1-\phi[\mathfrak{M}_{AB}f_1]\,\big|\,|B|\\ &\quad+\lim_{\delta\to 0}\ \sup_{F\in\Gamma_\delta(I)}\ \sum_{B\in F}\big|\,\mathfrak{M}_B\phi:f_2-\phi[\mathfrak{M}_{AB}f_2]\,\big|\,|B|=0,\end{aligned}$$

since ϕ is dominated by the convex function ϕ_i with

$$\int_I\phi_i:f_i<\sum_{n=1}^{\infty}2^{(2n)^2}/2^{(2n)^2+n}=1\qquad (i=1, 2).$$

We shall now show that if ψ is any convex function dominating ϕ on $(-\infty, \infty)$, $\int_I\psi:f$ does not exist (finite). Let n be odd and $\geqq 1$ and consider the interval $[2n, 2n+2]$. On this closed interval ϕ_1 and ϕ_2 are linear, with derivatives that we may designate, respectively, by μ_1 and μ_2 $(\mu_1>\mu_2)$, and their graphs intersect at a point (α, β), $2n<\alpha<2n+2$. On the subintervals

$[2n, \alpha]$ and $[\alpha, 2n+2]$ we have respectively $\phi=\phi_1$ and $\phi=\phi_2$. The function ζ defined thus on $(-\infty, \infty)$,

$$\zeta(y) = \begin{cases} \phi_1(y) & \text{for } 2n \leqq y \leqq \alpha, \\ \text{linear} & \text{for } -\infty < y \leqq \alpha, \\ \phi_2(y) & \text{for } \alpha \leqq y \leqq 2n+2, \\ \text{linear} & \text{for } \alpha \leqq y < \infty, \end{cases}$$

is concave on $(-\infty, \infty)$, with $\zeta(y) \leqq \psi(y)$ for $-\infty < y < \infty$. Hence there exists a function ϕ^* linear on $[2n, 2n+2]$, with derivative μ_0 and with $\phi^*(\alpha)=\beta$, satisfying the conditions

$$\mu_1 \geqq \mu_0 \geqq \mu_2,$$
$$\phi^*(y) \leqq \psi(y) \qquad \text{for } 2n \leqq y \leqq 2n+2.$$

Now a line through the point (α, β) with slope μ cuts the verticals through $(2n, 0)$ and $(2n+2, 0)$, respectively, in points whose ordinates are

$$h_1(\mu) = \beta + \mu(2n - \alpha), \qquad h_2(\mu) = \beta + \mu(2n + 2 - \alpha).$$

From the condition $\psi(y) \geqq \phi(y) \geqq 0$ on $-\infty < y < \infty$, we infer

$$\begin{aligned} \int_I \psi : f &> \int_{I_n} \psi : f + \int_{I_{n+1}} \psi : f = \psi(2n)/2^{(2n)^2+n} + \psi(2n+2)/2^{(2n+2)^2+n+1} \\ &\geqq \phi^*(2n)/2^{(2n)^2+n} + \phi^*(2n+2)/2^{(2n+2)^2+n+1} \\ &= h_1(\mu_0)/2^{(2n)^2+n} + h_2(\mu_0)/2^{(2n+2)^2+n+1} = H(\mu_0), \end{aligned}$$

where

$$H(\mu) = h_1(\mu)/2^{(2n)^2+n} + h_2(\mu)/2^{(2n+2)^2+n+1}.$$

Since H is a linear function of μ, at least one of the inequalities

$$H(\mu_0) \geqq H(\mu_1), \qquad H(\mu_0) \geqq H(\mu_2)$$

holds. But, since no one of the numbers $h_i(\mu_j)$ $(i, j=1, 2)$ can be negative, we have

$$\begin{aligned} H(\mu_1) &\geqq h_2(\mu_1)/2^{(2n+2)^2+n+1} = 2^{(2n+3)^2}/2^{(2n+2)^2+n+1} = 2^{3n+4}, \\ H(\mu_2) &\geqq h_1(\mu_2)/2^{(2n)^2+n} = 2^{(2n+1)^2}/2^{(2n)^2+n} = 2^{3n+1}, \end{aligned}$$

whence

$$\int_I \psi : f > 2^{3n+1}$$

and the fact that $\int_I \psi : f$ does not exist (finite) follows from the arbitrariness of the odd integer n.

Another question which it is natural to raise is answered by

EXAMPLE B. The following example shows that in Theorem (5.13) the hypothesis that $|\phi|$ be dominated by a *convex* function ψ cannot be weakened to the hypothesis that $|\phi|$ be dominated by a *strictly monotone* function ψ even if one takes $E=I$ and insists that f be continuous on I save at one point, that no sampling be indulged in, and that the restriction $F\in\Gamma_3(I)$ be imposed.

Our example is one in which f is non-negative on I and ϕ itself is positive, continuous, and strictly continuous on the interval $0\leqq y<\infty$, with $\int_I\phi:f<\infty$. In order that our purposes later in §11 may also be served we construct this example to satisfy the additional condition

$$\lim_{y\to\infty}\phi(y)/y=\infty.$$

To fit completely into the situation of §11 we should extend the interval on which ϕ is defined to $(-\infty,\infty)$ by setting

$$\phi(y)=16+y^2 \qquad \text{for } -\infty<y<0,$$

for example, so that ϕ is continuous and positive on $(-\infty,\infty)$ with $\lim_{|y|\to\infty}\phi(y)/|y|=\infty$.

Let

$$a_n=1/2^{2^{2n}},\qquad m_n=(a_{n+1})^{1/2}\qquad (n=1,2,3,\cdots),$$

so that $\lim_{n\to\infty}a_n=0$ and for each n we have

$$a_{n+1}=a_n^4,\qquad m_n^2=a_{n+1},\qquad m_n=a_n^2,\qquad a_{n+1}<m_n<a_n\leqq a_1=1/16.$$

On I let g be defined by the conditions

$$g(0)=0,\qquad g(x)=1\qquad \text{for } 1/16<x\leqq 1,$$

$$g(x)=\begin{cases} m_n^{-5/4} & \text{for } a_{n+1}<x\leqq m_n,\\ 1 & \text{for } m_n<x\leqq a_n,\end{cases}\qquad (n=1,2,3,\cdots),$$

and let

$$f(0)=0,\qquad f(x)=\int_x^1 g\qquad \text{for } 0<x\leqq 1.$$

The function f is then absolutely continuous on every interval $[\epsilon, 1]$, $\epsilon>0$, but the relation

$$\begin{aligned}\int_{a_{n+1}}^{m_n} g &= m_n^{-5/4}(m_n-a_{n+1})=m_n^{-5/4}(m_n-m_n^2)\\ &= m_n^{-1/4}(1-m_n)>(15/16)m_n^{-1/4}\qquad (n=1,2,3,\cdots)\end{aligned}$$

implies $\lim_{x\to 0}f(x)=\infty$.

The inequality

$$\text{(a)} \qquad \int_{a_{n+1}}^{a_n} f/(a_n - a_{n+1}) > f(m_n) \qquad (n = 1, 2, 3, \cdots)$$

may be shown as follows. For arbitrary n let

$$a = a_{n+1}, \qquad c = m_n, \qquad b = a_n.$$

Then we have

$$f(x) = f(b) + \int_x^b g = f(b) + (b - x) \qquad \text{for } c \leqq x \leqq b,$$

$$f(c) = f(b) + (b - c),$$

whence

$$\int_c^b f = (b - c)f(b) + (b - c)^2/2 = (b - c)f(c) - (b - c)^2/2;$$

and

$$f(x) = f(c) + \int_x^c g = f(c) + c^{-5/4}(c - x) \qquad \text{for } a \leqq x \leqq c,$$

$$\int_a^c f = (c - a)f(c) + c^{-5/4}(c - a)^2/2.$$

Thus we obtain

$$\int_a^b f = (b - a)f(c) + [c^{-5/4}(c - a)^2 - (b - c)^2]/2,$$

where

$$\begin{aligned} c^{-5/4}(c - a)^2 - (b - c)^2 &= c^{-5/4}(c - c^2)^2 - (b - c)^2 > c^{3/4}(1 - c)^2 - b^2 \\ &= c^{3/4}(1 - c)^2 - c > c^{3/4}(15/16)^2 - c \\ &= c^{3/4}[(15/16)^2 - c^{1/4}] \\ &> c^{3/4}[(15/16)^2 - (1/16)^{1/4}] > 0. \end{aligned}$$

From the definition of g we have

$$0 < g(x) \leqq x^{-5/4} \qquad \text{for } 0 < x \leqq 1,$$

whence

$$\text{(b)} \qquad 0 \leqq f(x) = \int_x^1 g \leqq \int_x^1 t^{-5/4}dt < 4x^{-1/4} \qquad \text{for } 0 < x \leqq 1.$$

Thus we infer $f \in L(I)$. Moreover, as x decreases from 1 toward 0, f is strictly increasing as well as continuous, and f increases from 0 toward ∞; that is, to each y, $0 \leqq y < \infty$, corresponds one and only one x with $0 < x \leqq 1$, $y = f(x)$. For $0 \leqq y < \infty$ we may therefore define $\phi(y)$ by the conditions

$$\phi(0) = 16, \quad \phi[f(x)] = 16x^{-1/2} \qquad \text{for } 0 < x \leqq 1.$$

Clearly, ϕ is continuous and strictly increasing on the infinite interval $0 \leqq y < \infty$, with $\int_I \phi : f < \infty$. On the other hand, from (a) we have for each n

$$\begin{aligned}(8.1) \qquad \phi\left[\int_{a_{n+1}}^{a_n} f/(a_n - a_{n+1})\right](a_n - a_{n+1}) &> \phi[f(m_n)](a_n - a_{n+1}) \\ &= 16m_n^{-1/2}(a_n - a_{n+1}) = 16a_n^{-1}(a_n - a_n^4) \\ &= 16(1 - a_n^3) > 16(1 - a_n) \geqq 16(15/16) = 15,\end{aligned}$$

which implies

$$\lim_{\delta\to 0} \sup_{F \in \Gamma_\delta(I)} \sum_{B \in F} \phi\left[\int_B f/|B|\right]|B| = \infty;$$

and from (b)

$$\phi[f(x)] = 16x^{-1/2} = (4x^{-1/4})^2 > [f(x)]^2 \qquad \text{for } 0 < x \leqq 1,$$

which implies

$$\lim_{y\to\infty} \phi(y)/y = \infty.$$

Finally, we may point out that no real generality is added to Theorem (5.13), when diam $E < \infty$ and $\phi \in C((-\infty, \infty))$, by assuming only deferred application of the dominance of $|\phi|$ by ψ or of the convexity of the dominant function ψ. For, if the conditions

$$\phi \in C((-\infty, \infty)), \qquad |\phi(y)| \leqq \psi(y) \qquad \text{for } 0 < M \leqq |y| < \infty,$$

$$\psi \text{ convex on } (-\infty, \infty), \qquad \int_E \psi : f < \infty$$

are satisfied, let $\alpha = \max_{|y| \leqq M} |\phi(y)|$. Then $\psi_1(y) = \psi(y) + \alpha$ is convex on $(-\infty, \infty)$, $|\phi|$ is $\leqq \psi_1$ on $(-\infty, \infty)$ and $\int_E \psi_1 : f = \int_E \psi : f + \int_E \alpha < \infty$. Again, if the conditions

$$\phi \in C((-\infty, \infty)), \quad |\phi(y)| \leqq \psi(y) \text{ on } (-\infty, \infty), \quad \psi \in C((-\infty, \infty)),$$

$$\psi \text{ convex for } -\infty < y \leqq M \text{ and for } M \leqq y < \infty, \qquad \int_E \psi : f < \infty$$

are fulfilled, let $\alpha = \max_{|y| \leqq M} |\psi(y)|$. Then $\psi_1(y) = \max [\psi(y), \alpha]$ is convex, and $|\phi|$ is $\leqq \psi_1$, on $(-\infty, \infty)$. If $\zeta(y) = \psi_1(y) - \psi(y)$ on $(-\infty, \infty)$, it is easily seen that $\zeta \in UC((-\infty, \infty))$, whence $f \in L_\zeta(E)$ by Theorem (5.2); thus $\int_E \psi_1 : f = \int_E \psi : f - \int_E \zeta : f$ exists.

9. **The essential nature of the hypothesis of continuity on ϕ.** The extent to which this hypothesis is essential to the result embodied in Theorem (5.12) is rather clearly indicated by the following theorem.

(9.1) THEOREM. *If ϕ is a function on $(-\infty, \infty)$ and $f \in BL_\phi(I)$ implies*

$$\lim_{n\to\infty} (1/2^n)\sum_{j=1}^{2^n} \phi\left[2^n \int_{(j-1)/2^n}^{j/2^n} f\right] = \int_I \phi : f,$$

then $\phi(y)$ is continuous on $(-\infty, \infty)$.

To prove this theorem we shall show that the assumption of the hypothesis and the contrary of the conclusion leads to a contradiction. For simplicity we first consider the case in which $\phi(y)$ has an hypothetical discontinuity at $y=0$, with

$$\text{(a)} \qquad \phi(-1) = \phi(0) = \phi(1) = 0, \qquad \limsup_{y\to 0+} \phi(y) > \alpha > 0.$$

We make these assumptions, and with the object of constructing a function $f \in BL_\phi(I)$ which will yield a contradiction we proceed presently to establish a lemma. For this purpose the following notation will be convenient:

$$\{j, n\} = \mathop{E}_{x} [(j-1)/2^n \leqq x \leqq j/2^n] \quad (n = 1, 2, 3, \cdots; j = 1, 2, \cdots, 2^n),$$

$$S_n(f) = (1/2^n)\sum_{j=1}^{2^n} \phi\left[2^n \int_{\{j,n\}} f\right],$$

$$P(f) = \mathop{E}_{x} [\,|f(x)| > 0].$$

We also define certain sequences of functions as follows. For $n \geqq 1$, $0<\xi<1$, and $A \subset I$ let $h_{n,\xi}$ be defined on I by

$$h_{n,\xi}(x) = \begin{cases} 1 & \text{for } j/2^n - \xi/2^n \leqq x \leqq j/2^n \qquad (j = 1, 2, \cdots, 2^{n-1}), \\ -1 & \text{for } 3/4 - \xi/4 \leqq x \leqq 3/4 + \xi/4, \\ 0 & \text{otherwise;} \end{cases}$$

let $H_{n,\xi}$ be defined on $(-\infty, \infty)$ to be periodic of period $1/2^n$ with

$$H_{n,\xi}(x/2^n) = h_{n,\xi}(x) \qquad \text{for } x \in I;$$

let $\delta_{n,A}$ be defined on I by

$$\delta_{n,A}(x) = \begin{cases} 0 & \text{for } x \in \{j, n\} \text{ if } A\{j, n\} \neq 0 \qquad (j = 1, 2, \cdots, 2^n), \\ 1 & \text{otherwise;} \end{cases}$$

and let $g_{n,\xi,A}$ be defined on I by

$$g_{n,\xi,A}(x) = \delta_{n,A}(x) H_{n,\xi}(x) \qquad \text{for } x \in I,$$

so that $P(g_{n,\xi,A})$ is a closed set. These functions being defined, we assert

LEMMA 1. *If A is a closed subset of I and $0<\xi<1$ we have*
(i) $g_{n,\xi,A}(x)=0$ *for* $x \in \{j, n\}$ *if* $A\{j, n\} \neq 0$ $(n=1, 2, 3, \cdots; j=1, 2, \cdots, 2^n)$;

(ii) *for integers* n_1, n_2 *with* $n_2 \geqq n_1 \geqq 1$,

$$\int_{[j,n_1]} g_{n_2,\xi,A}(x)dx = 0 \qquad (j = 1, 2, \cdots, 2^{n_1});$$

(iii) (*the hypothesis of Theorem* 8 *being assumed*)

$$\lim_{n\to\infty} S_{2n}(g_{n,\xi,A}) = (1 - |A|)\phi(\xi)/2.$$

Conclusion (i) is immediate from the definition of $\delta_{n,A}$. To prove the other conclusions we regard A and ξ as fixed, drop these letters as subscripts, set

$$\delta_n([(j-1)/2^n + j/2^n]/2) = \mu_{n,j} \qquad (n = 1, 2, 3, \cdots; j = 1, 2, \cdots, 2^n),$$

and proceed as follows. From the definition of h_n we have

$$\int_I h_n = 0 \qquad (n = 1, 2, 3, \cdots);$$

whence, for such n and $j = 1, 2, \cdots, 2^n$,

$$\int_{[j,n]} g_n(x)dx = \mu_{n,j}\int_{[j,n]} H_n(x)dx = \mu_{n,j}\int_{[1,n]} H_n(x)dx$$

$$= \frac{\mu_{n,j}}{2^n}\int_I H_n(x/2^n)dx = \frac{\mu_{n,j}}{2^n}\int_I h_n(x)dx = 0.$$

For $n_2 \geqq n_1 \geqq 1$ we then have

$$\int_{[j,n_1]} g_{n_2} = \sum_{\nu=(j-1)2^{n_2-n_1}+1}^{j2^{n_2-n_1}} \int_{[\nu,n_2]} g_{n_2} = 0 \qquad (j = 1, 2, \cdots, 2^{n_1}),$$

and (ii) is proved.

As for (iii) we observe first

$$\lim_{n\to\infty} S_n(h_n) = \lim_{n\to\infty} (1/2^n)\sum_{j=1}^{2^n} \phi\left[2^n \int_{[j,n]} h_n\right]$$

$$\text{(b)} \qquad = \lim_{n\to\infty} (1/2^n)\left\{\sum_{j=1}^{2^{n-1}} + \sum_{j=2^{n-1}+1}^{2^n}\right\}\phi\left[2^n\int_{[j,n]} h_n\right]$$

$$= (1/2^n)\phi(\xi)2^{n-1} + \int_{1/2}^1 \phi : h_n = \phi(\xi)/2,$$

by applying the hypothesis of Theorem (9.1) to the function on I which is zero for $x \in [0, 1/2]$ and equal to h_n for $x \in [1/2, 1]$. Next we note, for $\nu = 1, 2, \cdots, 2^n$,

$$\sum_{j=(\nu-1)2^n+1}^{\nu 2^n} \phi\left[2^{2n}\int_{(j,2n)} g_n(x)dx\right] = \sum_{j=(\nu-1)2^n+1}^{\nu 2^n} \phi\left[2^{2n}\int_{(j,2n)} \delta_n(x)H_n(x)dx\right]$$

$$= \mu_{n,\nu} \sum_{j=(\nu-1)2^n+1}^{\nu 2^n} \phi\left[2^{2n}\int_{(j,2n)} H_n(x)dx\right] = \mu_{n,\nu}\sum_{j=1}^{2^n} \phi\left[2^{2n}\int_{(j,2n)} H_n(x)dx\right]$$

$$= \mu_{n,\nu}\sum_{j=1}^{2^n} \phi\left[2^{n}\int_{(j,n)} H_n(x/2^n)dx\right] = \mu_{n,\nu}\sum_{j=1}^{2^n} \phi\left[2^{n}\int_{(j,n)} h_n(x)dx\right]$$

$$= \mu_{n,\nu}2^nS_n(h_n).$$

Hence for integral $n \geqq 1$, we have

$$S_{2n}(g_n) = (1/2^{2n})\sum_{j=1}^{2^{2n}} \phi\left[2^{2n}\int_{(j,2n)} g_n(x)dx\right]$$

$$= (1/2^{2n})\sum_{\nu=1}^{2^n}\sum_{j=(\nu-1)2^n+1}^{\nu 2^n} \phi\left[2^{2n}\int_{(j,2n)} g_n(x)dx\right]$$

$$= (1/2^{2n})\sum_{\nu=1}^{2^n} \mu_{n,\nu}2^nS_n(h_n) = S_n(h_n)\sum_{\nu=1}^{2^n}\mu_{n,\nu}/2^n = S_n(h_n)\int_I \delta_n(x)dx.$$

From the fact that $I-A$ is an open set, the density of the set of points $j/2^n$ ($n=1, 2, 3, \cdots$; $j=0, 1, \cdots, 2^n$) in I, and the definition of δ_n one may infer $\lim_{n\to\infty}\int_I\delta_n(x)dx = 1-|A|$. Using also (b) one obtains conclusion (iii).

Returning now to the proof of Theorem (9.1), let η be an arbitrary number satisfying the inequality $0<\eta<1$, and for each integer $n>0$ let ξ_n be a number satisfying the conditions $0<\xi_n<\eta/2^n$, $\phi(\xi_n)>\alpha$.

We are ready to begin an inductive definition of a sequence of functions f_n ($n=1, 2, 3, \cdots$). Let A_0 be the vacuous set; and let $N_1 \geqq 1$ be so large that, upon defining $f_1 = g_{N_1,\xi_1,A_0}$, we have in accordance with Lemma 1 (iii),

$$S_{2N_1}(f_1) \geqq \phi(\xi_1)/2 - 1/2.$$

For the second stage of the procedure let $A_1 = P(f_1)$; and let $N_2 > 2N_1$ be so large that, upon defining $f_2 = g_{N_2,\xi_2,A_1}$, we have in accordance with Lemma 1 (iii) and the hypothesis of Theorem (9.1),

$$S_{2N_2}(f_2) \geqq (1-|A_1|)\phi(\xi_2)/2 - 1/2^2,$$
$$|S_{2N_2}(f_1)| \leqq 1/2^2.$$

We observe $|A_1| = \xi_1$.

For the third stage, let $A_2 = P(f_1)+P(f_2) = A_1+P(f_2)$; and let $N_3 > 2N_2$ be so large that, upon defining $f_3 = g_{N_3,\xi_3,A_2}$, we have in accordance with Lemma 1 (iii) and the hypothesis of Theorem (9.1),

$$S_{2N_3}(f_3) \geqq (1-|A_2|)\phi(\xi_3)/2 - 1/2^3,$$
$$|S_{2N_3}(f_1+f_2)| \leqq 1/2^3.$$

The inequality $|A_2| < \xi_1+\xi_2$ should be noted. In general, f_j $(j=1, 2, \cdots, n-1)$ having been defined, let $A_{n-1}=A_{n-2}+P(f_{n-1})$; and let $N_n>2N_{n-1}$ be so large that upon defining $f_n=f_{N_n,\xi_n,A_{n-1}}$, we have

$$\text{(c)}\qquad \begin{aligned} &S_{2N_n}(f_n) \geqq (1 - |A_{n-1}|)\phi(\xi_n)/2 - 1/2^n, \\ &|S_{2N_n}(f_1 + f_2 + \cdots + f_{n-1})| \leqq 1/2^n. \end{aligned}$$

We note $|A_{n-1}| < \xi_1+\xi_2+ \cdots +\xi_{n-1}$.

Finally we define

$$F(x) = \sum_{j=1}^{\infty} f_j(x) \quad \text{for } x \in I, \qquad A = \sum_{j=1}^{\infty} A_j,$$

observing that $P(f)=A$ and that $A_j \subset A_{j+1}$ $(j=1, 2, 3, \cdots)$ implies

$$|A| = \lim_{j\to\infty} |A_j| \leqq \sum_{j=1}^{\infty} \xi_j < \eta.$$

For $n \geqq 2$ we may write

$$F = \sum_{\nu=1}^{n-1} f_\nu + f_n + \sum_{\nu=n+1}^{\infty} f_\nu$$

and obtain

$$\text{(d)}\qquad S_{2N_n}(F) = S_{2N_n}\left(\sum_{\nu=1}^{n-1} f_\nu\right) + S_{2N_n}(f_n).$$

This arises from the fact that for each j $(j=1, 2, \cdots, 2^{2N_n})$ we have by Lemma 1 (ii)

$$\int_{(j,2N_n)} \left(\sum_{\nu=n+1}^{\infty} f_\nu\right) = \sum_{\nu=n+1}^{\infty} \int_{(j,2N_n)} f_\nu = 0,$$

and by Lemma 1 (i)

$$\left[\int_{(j,2N_n)} \left(\sum_{\nu=1}^{n-1} f_\nu\right)\right]\left[\int_{(j,2N_n)} f_n\right] = 0,$$

which together imply

$$\begin{aligned} \phi\left[2^{2N_n}\int_{(j,2N_n)} F\right] &= \phi\left[2^{2N_n}\int_{(j,2N_n)} \left(\sum_{\nu=1}^{n-1} f_\nu\right) + 2^{2N_n}\int_{(j,2N_n)} f_n\right] \\ &= \phi\left[2^{2N_n}\int_{(j,2N_n)} \left(\sum_{\nu=1}^{n-1} f_\nu\right)\right] + \phi\left[2^{2N_n}\int_{(j,2N_n)} f_n\right]. \end{aligned}$$

From (c) and (d) we have

$$S_{2N_n}(F) \geqq (1 - |A_{n-1}|)\phi(\xi_n)/2 - 1/2^{n-1} \geqq (1 - |A_{n-1}|)\alpha/2 - 1/2^{n-1},$$

whence

$$\lim_{n\to\infty} S_{2N_n}(F) \geqq (1 - |A|)\alpha/2 > (1-\eta)\alpha/2.$$

But from the hypothesis of Theorem (9.1) we obtain

$$\lim_{n\to\infty} S_n(F) = \int_I \phi:F = 0,$$

an obvious contradiction. This completes the proof of Theorem (9.1) in the case characterized by conditions (a).

We shall now show how a proof of Theorem (9.1) in the general case can be obtained from the particular case already considered. Let β stand for the class of functions f each of which is measurable on I and assumes no more than three values. A function ϕ defined on $(-\infty, \infty)$ will be said to possess *property* (B) if and only if

$$f \in \beta \text{ implies } \lim_{n\to\infty} (1/2^n)\sum_n \phi\left[2^n \int_{[j,n]} f\right] = \int_I \phi:f.$$

The result obtained above may now be formulated thus:

(i) *If* ϕ *has property* (B), *with* $\phi(-1)=\phi(0)=\phi(1)=0$, *then* $\limsup_{y\to 0+}\phi(y)$ *is* $\leqq 0$.

We assert

(ii) *A necessary and sufficient condition that* $\phi(y)$ *be continuous for all* y *is that it possess property* (B).

The necessity of the condition may be inferred at once from Theorem (5.2). To prove the sufficiency we first verify without difficulty that if ϕ_1 and ϕ_2 have property (B), so also do $\phi_1+\phi_2$ and $c\phi_1(ay+b)$, where a, b, and c are arbitrary real numbers. Now let ϕ be assumed to have property (B) and to be discontinuous at an arbitrary point y_1; then $\phi_1(y)=\phi(y-y_1)$ has property (B) and is discontinuous at $y=0$. Also, for $-\infty<y<\infty$, let

$$\phi_2(y) = \phi_1(-1)y(1-y)/2 + \phi_1(0)(y+1)(y-1) + \phi_1(1)(-y)(1+y)/2,$$

so that ϕ_2, being continuous, has property (B). The function $\phi_3=\phi_1+\phi_2$ then has property (B), with $\phi_3(-1)=\phi_3(0)=\phi_3(1)=0$, and ϕ_3 is discontinuous at $y=0$. Of the four functions $\pm\phi_3(\pm y)$, at least one is a function ϕ_4 possessing property (B), with $\phi_4(-1)=\phi_4(0)=\phi_4(1)=0$ and with $\limsup_{y\to 0+}\phi_4(y)>0$, in contradiction to (i).

This completes the proof of Theorem (9.1).

REMARK. In connection with Theorem (9.1) it is worth noting that the conditions $\phi\in C((-\infty, \infty))$, $f\in BL_\phi(I)$ imply the existence of a function ψ convex on $(-\infty, \infty)$ and depending upon f, with $|\phi(y)| \leqq \psi(y)$ on $(-\infty, \infty)$ and $f\in L_\psi(I)$. It would be of some interest to prove or disprove either of the following conjectures. (1) The hypothesis of Theorem (9.1) implies that to

each $f\in L_\phi(I)$ corresponds a function ψ convex on $(-\infty, \infty)$, with $|\phi(y)|\leqq\psi(y)$ on $(-\infty, \infty)$ and $f\in L_\psi(I)$. (2) The hypothesis of Theorem (9.1) implies the existence of a function ψ convex on $(-\infty, \infty)$, with $|\phi(y)|\leqq\psi(y)$ on $(-\infty, \infty)$ and $L_\phi(I)\subset L_\psi(I)$ (that is, $L_\phi(I)=L_\psi(I)$). Attempts which we have made in this direction have met with no success.

10. **Concerning the replacement of hypothesis** (1.6) **by** (2.10). In this section we propose to determine as definitely as possible under what circumstances the sample sets need only be of positive measure; that is, the "k-hypothesis" (1.6) can be relaxed to (2.10). A basic, though simple, result of this kind is embodied in Theorem (5.1). In RS it has been seen that, in case $E=I$ and $\phi(y)=y$ for all y, the question turns entirely on whether $f\in R(I)$, the class of functions Riemann integrable on I. Thus it is natural to restrict the considerations of this section to the case in which E is a bounded interval, say $E=I$.

We begin by establishing

(10.1) THEOREM. *If $f\in L(I)$, $\phi\in C(\mathfrak{R}(f))$, $\phi:f\in L(I)-R(I)$, and* (1.6) *is replaced by* (2.10),

$$\lim_{\delta\to 0}\sum_{B\in F\in\Gamma_\delta(I)}\phi[\mathfrak{M}_{AB}f]\,|\,B\,|$$

does not exist. Thus, when $\phi:f\in L(I)-R(I)$, the k-hypothesis is indispensable to Theorems (5.2)–(5.13).

Proof. However small $\delta>0$ may be, let $F\in\overline{\Gamma}_\delta(I)$. If $\phi:f$ is not essentially bounded on I, there is at least one interval $B\in F$ on which $\phi:f$ is not essentially bounded; that is, for arbitrary $M>0$ there exists a measurable set $A\subset B$ with

$$|\,A\,|>0,\qquad |\,\phi[f(x)]\,|>M\qquad\text{for } x\in A.$$

The set

$$C = B\mathop{E}_{x}[f(x)\ \bar{\in}\ \mathfrak{R}(f)\text{ for } f \text{ on } B]$$

is of measure zero. Hence $A(B-C)$ is of positive measure, and there exist points x_1 with

$$|\,\phi[f(x_1)]\,|>M,\qquad x_1\in\mathfrak{R}(f)\qquad\text{for } f \text{ on } B;$$

that is, there exists a measurable set $A_1\subset B$ with

$$|\,A_1\,|>0,\qquad |\,\phi[\mathfrak{M}_{A_1}f]\,|>M.$$

Thus the sum in question is unbounded.

Secondly, if $\phi:f$ is essentially bounded on I and $B\in F$, let

$$\operatorname*{ess\,inf}_{x\in B}\phi[f(x)]=l,\qquad \operatorname*{ess\,sup}_{x\in B}\phi[f(x)]=L.$$

The same reasoning now suffices to show that the sum in question can be made arbitrarily close to either the essential lower or essential upper Darboux sum for $\phi:f$ corresponding to the family of intervals F. Hence the $\lim_{\delta\to 0}$ of the sum does not exist.

We next proceed in the direction of obtaining a favorable result when $f\in R(I)$. Without the k-hypothesis, even the conditions $f\in C(I)$ and $\phi\in C(\mathfrak{R}(f))$ are shown to be insufficient for the existence of $\int_I \phi:f$, and hence insufficient for (1.9), by the trivial example

$$f(x) = x \quad \text{for} \quad x \in I, \qquad \phi(y) = 1/y \quad \text{for} \quad y \in (0, 1) = \mathfrak{R}(f).$$

The nature of Theorems (5.8)–(5.12) suggests that we consider assuming

$$(10.2) \qquad f \in R(I), \quad \phi \in C(\mathfrak{R}(f)), \quad \lim_{\delta\to 0} \sup_{F\in\Gamma_\delta(I)} \sum_{B\in F} \left| \phi[\mathfrak{M}_{AB}f] \right| \left| B \right| < \infty.$$

We assert, however,

(10.3) Theorem. *The conditions* (10.2) *are equivalent to the conditions*

$$f \in R(I), \qquad \phi \in BC(\mathfrak{R}(f)).$$

Proof. It is evident that this pair of conditions implies the set (10.2). On the other hand, if $\phi\,\bar{\in}\,B(\mathfrak{R}(f))$, let $a\,\bar{\in}\,\mathfrak{R}(f)$ be an endpoint of $\mathfrak{R}(f)$ in every neighborhood of which ϕ is unbounded; and let F be an arbitrary family $\in\Gamma_\delta(I)$ for any $\delta>0$. Then at least one interval $B\in F$ has the property that the ess inf, or ess sup, of f on B is a. Since $a\,\bar{\in}\,\mathfrak{R}(f)$, it follows at once that as AB ranges over the subsets of B having measure >0, $\mathfrak{M}_{AB}f$ assumes all values in a unilateral neighborhood of a. Hence $\phi[\mathfrak{M}_{AB}f]|B|$ is an unbounded function of AB, and the sum in (10.2) is unbounded for each $\delta>0$.

As a preliminary step in the desired direction we next establish

(10.4) Theorem. *The conditions $f\in R(I)$ and $\phi\in C(\overline{\mathfrak{R}(f)})$ are sufficient for* (1.9), *even with* (1.6) *replaced by* (2.10).

Proof. We may assume f itself to be an element of $R(I)$, with $f(x)\in\overline{\mathfrak{R}(f)}$ for $x\in I$. $\overline{\mathfrak{R}(f)}$ is a bounded interval, whence $\phi\in UC(\overline{\mathfrak{R}(f)})$. Thus to $\epsilon>0$ corresponds $\eta=\eta(\epsilon)>0$ such that the conditions

$$y_1, y_2 \in \overline{\mathfrak{R}(f)}, \quad |y_1 - y_2| < \eta \quad \text{imply} \quad |\phi(y_1) - \phi(y_2)| < \epsilon/6;$$

and we then have

$$|\phi(y_1) - \phi(y_2)| < \epsilon/6 + \epsilon|y_1 - y_2|/(6\eta) \qquad \text{for } y_1, y_2 \in \overline{\mathfrak{R}(f)}.$$

Let g and h be functions continuous on I and satisfying the conditions

$$g(x) \leqq f(x) \leqq h(x) \quad \text{for} \quad x \in I, \qquad \min_{x\in I} g(x) \in \overline{\mathfrak{R}(f)}, \qquad \max_{x\in I} g(x) \in \overline{\mathfrak{R}(f)},$$

$$\min_{x\in I} h(x) \in \overline{\mathfrak{R}(f)}, \qquad \max_{x\in I} h(x) \in \overline{\mathfrak{R}(f)}, \qquad \int_I (h - g) < \eta.$$

Let $\delta_1=\delta_1(h-g, \phi_1, \eta)>0$ correspond to η for the functions $h-g$ and $\phi_1(y)=y$ $(-\infty<y<\infty)$ in accordance with Theorem (5.1); let $\delta_2=\delta_2(h, \phi, \epsilon/6)>0$ correspond to $\epsilon/6$ for the functions h and ϕ in accordance with Theorem (5.1); and let $\delta=\min[\delta_1, \delta_2]$. For $F\in\Gamma_\delta(I)$ we then have

$$\sum_{B\in F} |\mathfrak{M}_B\phi:f - \phi[\mathfrak{M}_{AB}f]|\,|B|$$
$$\leqq \sum_{B\in F}\int_B |\phi:f-\phi:h| + \sum_{B\in F} |\mathfrak{M}_B\phi:h - \phi[\mathfrak{M}_{AB}h]|\,|B|$$
$$+ \sum_{B\in F} |\phi[\mathfrak{M}_{AB}h] - \phi[\mathfrak{M}_{AB}f]|\,|B|.$$

On the right the first term is

$$\int_I |\phi:f-\phi:h| < \epsilon/6 + \epsilon\int_I |f-h|/(6\eta) \leqq \epsilon/6 + \int_I (h-g)/(6\eta)$$
$$< \epsilon/6 + \epsilon/6 = \epsilon/3.$$

By the choice of δ $(\leqq\delta_2)$ the second term on the right is $<\epsilon/6$. The third term is not greater than

$$\sum_{B\in F} \{\epsilon/6 + \epsilon|\mathfrak{M}_{AB}h - \mathfrak{M}_{AB}f|/(6\eta)\}\,|B|$$
$$\leqq \epsilon/6 + \epsilon/(6\eta)\sum_{B\in F}\left\{|B|\,\mathfrak{M}_{AB}(h-g) - \int_B (h-g) + \int_B (h-g)\right\}$$
$$\leqq \epsilon/6 + \epsilon/(6\eta)\sum_{B\in F} |\mathfrak{M}_{AB}(h-g) - \mathfrak{M}_B(h-g)|\,|B| + \epsilon/(6\eta)\int_I (h-g),$$

which by the choice of δ $(\leqq\delta_1)$ is less than

$$\epsilon/6 + [\epsilon/(6\eta)]\eta + \epsilon/6 = \epsilon/2.$$

The proof is now complete.

For use in demonstrating the next theorem the following lemma will be convenient.

(10.5) LEMMA. *Let* $f\in R(I)$, $\phi\in BC(\mathfrak{R}(f))$, $\sup_{y\in\mathfrak{R}(f)}|\phi(y)|<M<\infty$,

$$\alpha = \inf_{x\in I} f(x), \qquad \beta = \sup_{x\in I} f(x), \qquad \beta-\alpha>0,$$
$$\alpha\,\bar{\bar{\varepsilon}}\,\mathfrak{R}(f) \qquad [\beta\,\bar{\bar{\varepsilon}}\,\mathfrak{R}(f)],$$

$0<\eta<(\beta-\alpha)/2$, *and*

$$E_\eta = \underset{x}{E}\,[f(x)\leqq\alpha+\eta] \qquad [E_\eta = \underset{x}{E}\,[f(x)\geqq\beta-\eta]].$$

To $\epsilon>0$ *corresponds* $\eta=\eta(\epsilon)>0$ *and* $\delta=\delta(\eta)>0$ *such that if* $F\in\Gamma_\delta(I)$ *and* F_1

represents the subfamily of F *in which each set* B *has the property* $|BE_\eta|>0$, *we have* $\sum_{B\in F_1}|B|<\epsilon/M$, *whence*

$$\sum_{B\in F_1}|\mathfrak{M}_B\phi:f||B|<\epsilon \quad \textit{and} \quad \sum_{B\in F_1}|\phi[\mathfrak{M}_{AB}f]||B|<\epsilon.$$

Proof. We have $0=|E_x[f(x)=\alpha]|=\lim_{\eta\to 0}|E_\eta|$; let η satisfy the conditions

$$0<\eta<(\beta-\alpha)/2, \qquad |E_\eta|<\epsilon/M.$$

We may assume f itself to be an element of $R(I)$. Any point $x\in\overline{E}_\eta-E_\eta$, where $\overline{E}_\eta$ represents the closure of E_η, is then a point of discontinuity of f; and since the discontinuities of f are a set of measure zero, we have $|\overline{E}_\eta|=|E_\eta|<\epsilon/M$. Let

$$I-\overline{E}_\eta=\sum_i O_i,$$

where $\sum_i\{O_i\}$ is a countable set of disjoint open intervals. Let k be a finite number of these intervals with the property

$$\sum_{i=1}^{k}|O_i|>1-\epsilon/M.$$

From each end of each O_i $(i=1,2,\cdots,k)$ delete a semi-open interval to leave an open interval O_i', in such manner that we have

$$\sum_{i=1}^{k}|O_i'|>1-\epsilon/M;$$

and let the measure of the smallest deleted interval be δ $(\delta>0)$. For $F\in\Gamma_\delta(I)$ and $F_2=F-F_1$ we then have

$$\sum_{B\in F_2}|B|\geqq\sum_{i=1}^{k}|O_i'|>1-\epsilon/M, \qquad \sum_{B\in F_1}|B|<\epsilon/M,$$

$$\sum_{B\in F_1}|\mathfrak{M}_B\phi:f||B|<M\epsilon/M=\epsilon, \qquad \sum_{B\in F_1}|\phi[\mathfrak{M}_{AB}f]||B|\leqq M\epsilon/M=\epsilon.$$

(10.6) THEOREM. *The conditions* $f\in R(I)$ *and* $\phi\in BC(\mathfrak{R}(f))$ *are sufficient for* (1.9), *even with* (1.6) *replaced by* (2.10).

Proof. Let f itself be an element of $R(I)$. If $\mathfrak{R}(f)=\overline{\mathfrak{R}(f)}$, the conclusion follows from Theorem (10.4). If $\alpha=\inf_{x\in I}f(x)$, $\beta=\sup_{x\in I}f(x)$, $\beta-\alpha>0$, $\alpha\,\overline{\in}\,\mathfrak{R}(f)\,[\beta\,\overline{\in}\,\mathfrak{R}(f)]$, and $0<\eta<(\beta-\alpha)/2$, let

$$f_\eta(x)=\alpha+\eta \qquad \text{for } x\in \underset{x}{E}\,[f(x)<\alpha+\eta]$$

$$[f_\eta(x)=\beta-\eta \qquad \text{for } x\in \underset{x}{E}\,[f(x)>\beta-\eta]];$$

and for other values of x let $f_\eta(x)=f(x)$. One easily sees

$$f_\eta\in R(I), \qquad \mathfrak{R}(f_\eta)=\overline{\mathfrak{R}(f_\eta)}\subset\mathfrak{R}(f), \qquad \phi\in C(\overline{\mathfrak{R}(f_\eta)}).$$

Let η and $\delta_1>0$ correspond to $\epsilon/9$ in accordance with Lemma (10.5); let $\delta_2=\delta_2(f_\eta, \phi, \epsilon/9)>0$ correspond to $\epsilon/9$ for the functions f_η and ϕ in accordance with Theorem (10.4); and let $\delta=\min[\delta_1, \delta_2]$. For $F\in\Gamma_\delta(I)$ we then have

$$\begin{aligned}
&\sum_{B\in F} \left| \mathfrak{M}_B\phi:f - \phi[\mathfrak{M}_{AB}f]\right| \left| B \right| \\
&\quad \leqq \sum_{B\in F} \left| \mathfrak{M}_B\phi:f - \mathfrak{M}_B\phi:f_\eta\right| \left| B\right| + \sum_{B\in F} \left|\mathfrak{M}_B\phi:f_\eta - \phi[\mathfrak{M}_{AB}f_\eta]\right| \left| B\right| \\
&\qquad + \sum_{B\in F} \left|\phi[\mathfrak{M}_{AB}f_\eta] - \phi[\mathfrak{M}_{AB}f]\right| \left| B\right| \\
&\quad < \sum_{B\in F_1} \left|\mathfrak{M}_B\phi:f\right| \left| B\right| + \sum_{B\in F_1} \left|\mathfrak{M}_B\phi:f_\eta\right| \left| B\right| + \epsilon/9 \\
&\qquad + \sum_{B\in F_1} \left|\phi[\mathfrak{M}_{AB}f_\eta]\right| \left| B\right| + \sum_{B\in F_1} \left|\phi[\mathfrak{M}_{AB}f]\right| \left| B\right| < \epsilon,
\end{aligned}$$

where F_1 has the significance attached to it in Lemma (10.5).

The question of whether the k-hypothesis can be relaxed under the conditions

$$f \in L(I) - R(I), \qquad \phi:f \in R(I), \qquad \phi \in UC(\overline{\mathfrak{R}(f)})$$

is still open. That it sometimes can is illustrated by the trivial example

$$f \in L(I), \qquad \phi(y) = \text{const.} \qquad \text{for } -\infty < y < \infty.$$

That it sometimes cannot, even if ϕ is a very simple convex function and f is bounded, may be shown as follows. Let $\phi(y)=|y|$ for $-1\leqq y\leqq 1$; let $E\subset I$ be a non-dense perfect set, with $|E|=1/2$, and let $f(x)=1$ for $x\in E$, $f(x)=-1$ for $x\in I-E$. For each m $(m=1, 2, 3, \cdots)$ let $F_m\in\Gamma_1(I)$ be the family of 2^m equal subintervals of I. In each one of these intervals there exists a subset of $I-E$ of measure >0. In each one of at least half of these intervals there exists a subset of E of measure >0, and hence a subset AB with $|AB|>0$, such that $\mathfrak{M}_{AB}f=\int_{AB}f=0$. In any one of the remaining subintervals the maximum absolute value of the integral mean of f on any subset is 1. Hence we have

$$\liminf_{m\to\infty} \sum_{B\in F_m} \phi[\mathfrak{M}_{AB}f]\left| B\right| \leqq 1/2 < 1 = \int_I \phi:f.$$

This example being in hand, some interest may attach to

(10.7) THEOREM. *The conditions* $f\in L(I)$, $\phi:f\in R(I)$, ϕ *convex on* $\mathfrak{R}(f)$, *and* (1.6) *replaced by* (2.10), *imply*

$$\lim_{\delta\to 0} \sup_{F\in\Gamma_\delta(I)} \phi[\mathfrak{M}_{AB}f]\left| B\right| \leqq \int_I \phi:f < \infty.$$

Proof. We may assume $\phi:f$ itself to be an element of $R(I)$. Let ϵ be an arbitrary positive number and let g be a function continuous on I and satisfying the conditions

$$\phi[f(x)] \leqq g(x) \quad \text{for} \quad x \in I, \qquad \int_I (g - \phi:f) < \epsilon/2.$$

In accordance with Theorem (5.1) let $\delta = \delta(g, \phi_1, \epsilon/2) > 0$ correspond to $\epsilon/2$ for the functions g and $\phi_1(y) = y$ $(-\infty < y < \infty)$. Using Jensen's inequality we then have, for $F \in \Gamma_\delta(I)$,

$$\sum_{B \in F} \phi[\mathfrak{M}_{AB}f] \, | B | \leqq \sum_{B \in F} [\mathfrak{M}_{AB}\phi:f] \, | B | \leqq \sum_{B \in F} [\mathfrak{M}_{AB}g] \, | B |$$
$$< \int_I g + \epsilon/2 < \int_I \phi:f + \epsilon.$$

In the example last cited above f is in the class $L(I) - R(I)$ because of the fact that, though bounded, it has discontinuities at a set of positive measure. This gives rise to the question: do there exist examples in which it is impossible to relax the k-hypothesis when $f \in L(I) - R(I)$ because f is unbounded, even though f is improperly Riemann integrable on I, $\phi:f \in R(I)$, and (1.9) holds with the k-hypothesis in force? Should no such example exist, it would probably be possible to enlarge the class of cases already determined by Theorem (10.6) in which the k-hypothesis can be relaxed by obtaining an affirmative answer to some such question as the following. Do the conditions

$$f \in R^*(I), \qquad \phi:f \in R(I), \qquad \phi \in UC(\overline{\mathfrak{R}(f)})$$

imply that the k-hypothesis can be relaxed?

The answer to this question, however, is negative, as one may see from the ensuing example. Let

$$\phi(y) = \begin{cases} 0 \quad \text{for} \quad y = 2(2^n - 1) & (n = 0, 1, 2, \cdots); \\ 2^n \quad \text{for} \quad y = 2(2^n - 1) - 2^{n-1} & (n = 1, 2, 3, \cdots); \\ \text{linear between consecutive values of } y \text{ for which } \phi \text{ is already defined}; \end{cases}$$

$$f(x) = \begin{cases} 0 \quad \text{for} \quad 1/2^2 < x \leqq 1, \quad x = 0; \\ 2(2^n - 1) \quad \text{for} \quad 1/[4(2^{n+1} - 1)^2] < x \leqq 1/[4(2^n - 1)^2] \\ \qquad\qquad\qquad\qquad\qquad (n = 1, 2, 3, \cdots). \end{cases}$$

Then we have $f \in R^*(I)$, since it is a non-negative function dominated on I by f_1 where $f_1(x) = 1/x^{1/2}$ for $0 < x \leqq 1$, $f_1(0) = 0$; $\phi:f \in R(I)$, because it vanishes identically on I; $\mathfrak{R}(f) = \overline{\mathfrak{R}(f)} = E_y[0 \leqq y < \infty]$; and $\phi \in UC(\overline{\mathfrak{R}(f)})$, for its graph consists of the upper sides of an infinite sequence of similar triangles all with bases on the y-axis. If now $F \in \Gamma_1(I)$ is the family of 2^n equal subintervals of I, the terms of the sum

$$\sum_{B \in F} \phi[\mathfrak{M}_{AB}f] \, | B |$$

are all non-negative and the value of the term containing the interval $B=[0, 1/2^n]$ can be made arbitrarily large by a suitable choice of A. Thus the $\lim_{\delta\to 0}$ of the $\sup_{F\in\Gamma_\delta(I)}$ of such sums is infinite.

In the foregoing example ϕ is unbounded. It is natural therefore to ask if the situation is the same when ϕ is bounded; in answer we have theorems as follows.

(10.8) THEOREM. *The conditions* $f\in R^*(I)$ *and* $\phi\in BUC(\mathfrak{R}(f))$ *are sufficient for* (1.9), *even with* (1.6) *replaced by* (2.10).

Proof. To $\epsilon>0$ corresponds $\eta=\eta(\epsilon)>0$ such that the conditions

$$y_1, y_2 \in \mathfrak{R}(f), \quad |y_1 - y_2| < \eta \qquad \text{imply } |\phi(y_1) - \phi(y_2)| < \epsilon/16,$$

whence

$$|\phi(y_1) - \phi(y_2)| < \epsilon/16 + \epsilon|y_1 - y_2|/(16\eta) \qquad \text{for } y_1, y_2 \in \mathfrak{R}(f).$$

Let M satisfy the condition $|\phi(y)|<M<\infty$ for $y\in\mathfrak{R}(f)$.

Let $D\subset I$ be the set of points defined by the condition $x\in D$ if and only if $\text{ess lim sup}_{x'\to x}|f(x')|=\infty$. Clearly D is closed, with $|D|=0$. Let the open set $I-D$ be constituted of the disjoint open intervals O_i $(i=1, 2, 3, \cdots)$, and let m be such that $\sum_{i=1}^m|O_i|>1-\epsilon/(8M)$. From each end of O_i $(i=1, 2, 3, \cdots, m)$ delete an open interval of length $\epsilon/(16Mm)$ and designate the remaining closed interval by F_i, so that $\sum_{i=1}^m|F_i|>1-\epsilon/(4M)$. On each interval F_i the function f is essentially bounded. Let F_N be the truncated function f specified in the definition of the condition $f\in R^*(I)$ as given in §2, N being chosen to satisfy the conditions

$$\int_I |f - f_N| < \eta, \qquad \operatorname*{ess\,sup}_{x\in F_i} |f(x)| < N < \infty \qquad (i = 1, 2, \cdots, m).$$

Finally, let $\delta_1=\delta_1(f, \phi, \epsilon/8)>0$ correspond to $\epsilon/8$ for the functions f_N and ϕ in accordance with Theorem (10.6); let $\delta_2=\epsilon/(16Mm)$; and let $\delta=\min[\delta_1, \delta_2]$. For convenience let F_i' represent the closed interval obtained from F_i $(i=1, 2, \cdots, m)$ by deleting from each end of F_i a semi-open interval of length $\epsilon/(16Mm)$, so that we have $\sum_{i=1}^m|F_i'|>1-3\epsilon/(8M)$ and $|I-\sum_{i=1}^m F_i'|=1-\sum_{i=1}^m|F_i'|<3\epsilon/(8M)$. For $F\in\Gamma_\delta(I)$ let F_1 represent the subfamily of F defined by the condition $B\in F_1$ if and only if $B\sum_{i=1}^m F_i'\neq 0$; and let $F_2=F-F_1$. For $F\in\Gamma_\delta(I)$ we than have

$$\begin{aligned}\sum_{B\in F} |\mathfrak{M}_B\phi:f - \phi[\mathfrak{M}_{AB}f]|\,|B| &\leq \int_I |\phi:f - \phi:f_N| + \sum_{B\in F} |\mathfrak{M}_B\phi:f_N - \phi[\mathfrak{M}_{AB}f_N]|\,|B| \\ &\quad + \sum_{B\in F} |\phi[\mathfrak{M}_{AB}f_N] - \phi[\mathfrak{M}_{AB}f]|\,|B|.\end{aligned}$$

On the right the first term is less than

$$\epsilon/16 + \epsilon \int_I |f - f_N| /(16\eta) < \epsilon/16 + \epsilon/16 = \epsilon/8.$$

By our choice of δ ($\leqq \delta_1$) the second term on the right is $<\epsilon/8$. The third term may be written

$$\left\{ \sum_{B\in F_1} + \sum_{B\in F_2} \right\} |\phi[\mathfrak{M}_{AB}f_N] - \phi[\mathfrak{M}_{AB}f]| \, |B|.$$

Since δ is $\leqq \delta_2$, each set $B\in F_1$ contains only points of $\sum_{i=1}^m F_i$; whence, by choice of N and f_N,

$$\sum_{B\in F_1} |\phi[\mathfrak{M}_{AB}f_N] - \phi[\mathfrak{M}_{AB}f]| \, |B| = 0.$$

As for the remaining sum, from the relation $\sum_{B\in F_2} B \subset I - \sum_{i=1}^m F_i'$ we infer

$$\sum_{B\in F_2} |\phi[\mathfrak{M}_{AB}f_N] - \phi[\mathfrak{M}_{AB}f]| \, |B| \leqq 2M3\epsilon/(8M) = 3\epsilon/4,$$

which completes the proof.

(10.9) THEOREM. *The conditions $f\in R^*(I)$ and $\phi\in BC(\mathfrak{R}(f))$ are sufficient for* (1.9), *even with* (1.6) *replaced by* (2.10).

Proof. Let M satisfy the condition $|\phi(y)| < M < \infty$ for $y\in\mathfrak{R}(f)$; and let ϕ_i ($i=1, 2, 3, \cdots$) be a normal approximating sequence for ϕ on $\mathfrak{R}(f)$, as defined in (5.5). Let sets O_i, F_i, and F_i' be defined as in the proof of Theorem (10.8). Let the positive integer N satisfy the conditions

$$2M|E_N| < \epsilon/8 \quad \text{where} \quad E_N = \underset{x}{E}\,[|f(x)| > N],$$

$$\operatorname*{ess\,sup}_{x\in F_i} |f(x)| < N \qquad (i = 1, 2, \cdots, m).$$

Let $\delta_1 = \delta_1(f, \phi_N, \epsilon/8) > 0$ correspond to $\epsilon/8$ for the functions f and $\phi_N \in BUC(\mathfrak{R}(f))$ in accordance with Theorem (10.8); let $\delta_2 = \epsilon/(16Mm)$; and let $\delta = \min[\delta_1, \delta_2]$. For $F\in\Gamma_\delta(I)$ and F_1, F_2 defined as in the proof of Theorem (10.8) we then have

$$\begin{aligned}
&\sum_{B\in F} |\mathfrak{M}_B\phi:f - \phi[\mathfrak{M}_{AB}f]| \, |B| \\
&\qquad \leqq \int_I |\phi:f - \phi_N:f| + \sum_{B\in F} |\mathfrak{M}_B\phi_N:f - \phi_N[\mathfrak{M}_{AB}f]| \, |B| \\
&\qquad\qquad + \sum_{B\in F} |\phi_N[\mathfrak{M}_{AB}f] - \phi[\mathfrak{M}_{AB}f]| \, |B| \\
&\qquad < \epsilon/8 + \epsilon/8 + \left\{ \sum_{B\in F_1} + \sum_{B\in F_2} \right\} |\phi_N[\mathfrak{M}_{AB}f] - \phi[\mathfrak{M}_{AB}f]| \, |B|.
\end{aligned}$$

By the choice of N, each term in $\sum_{B\in F_1}$ vanishes and the sum $\sum_{B\in F_2}$ is not greater than $2M3\epsilon/(8M)=3\epsilon/4$.

11. Generalizations of boundedness, continuity, bounded variation, and absolute continuity of a function. The principal object of this section is to sketch in a background behind certain well known properties of the L_p $(p\geqq 1)$ classes of functions, which will throw these properties into rather clear relief. For our present purposes E will be taken as the unit interval I, f as a function whose domain includes I, and ϕ as a *non-negative* function whose domain includes the set of numbers

$$[f(b)-f(a)]/(b-a),\qquad \text{where } 0\leqq a<b\leqq 1. \tag{11.1}$$

The results to be obtained will lose only a little of their significance if ϕ is regarded as a *fixed* function whose domain includes the interval $(-\infty,\infty)$, while the function f is left free to vary; and this case is probably the one of greatest interest.

(11.2) DEFINITION. *The function f shall be said to be ϕ-bounded on I if and only if there exists a number M with*

$$\phi\left[\frac{f(b)-f(a)}{b-a}\right](b-a)\leqq M<\infty \qquad \text{for } 0\leqq a<b\leqq 1.$$

(11.3) DEFINITION. *For $a\in I$ the function f shall be said to be ϕ-continuous at a if and only if the condition*

$$\text{(a)}\qquad \lim_{x\in I,x\to a}\phi\left[\frac{f(x)-f(a)}{x-a}\right]|x-a|=0$$

is satisfied; f shall be said to be ϕ-continuous on I if and only if f is ϕ-continuous at each $a\in I$, uniformly ϕ-continuous on I if and only if condition (a) *holds uniformly with respect to a for $a\in I$.*

(11.4) DEFINITION. *The ϕ-total variation of f on I is*

$$T_I^\phi(f)=\limsup_{\delta\to 0}\sum_{j=1}^{N}\phi\left[\frac{f(x_j)-f(x_{j-1})}{x_j-x_{j-1}}\right](x_j-x_{j-1})\leqq\infty,$$

where

$$0=x_0<x_1<\cdots<x_N=1,\qquad \delta=\max_{1\leqq j\leqq N}(x_j-x_{j-1}).$$

If and only if $T_I^\phi(f)$ is $<\infty$, f shall be said to be of ϕ-bounded variation on I.

(11.5) DEFINITION. *The function f shall be said to be ϕ-absolutely continuous on I if and only if to each $\epsilon>0$ corresponds a $\delta=\delta(\epsilon)>0$ such that, $[a_j, b_j]$ $(a_j<b_j;\, j=1, 2, \cdots, N)$ being any finite set of closed subintervals of I disjoint except perhaps for common endpoints, the inequality $\sum_j(b_j-a_j)<\delta$ implies*

$$\sum_j \phi\left[\frac{f(b_j) - f(a_j)}{b_j - a_j}\right](b_j - a_j) < \epsilon.$$

Several consequences of these definitions are plain. (I) The defining properties here employed reduce to those of ordinary boundedness, continuity, and so on when $\phi(y) = |y|$ for $-\infty < y < \infty$. (II) If f is ϕ-absolutely continuous on I, it is uniformly ϕ-continuous on I and of ϕ-bounded variation on I. (III) If ϕ is bounded on the set of numbers (11.1), f is ϕ-absolutely continuous on I; in particular this hypothesis will be satisfied if f satisfies a Lipschitz condition (of order 1) on I with Lipschitz modulus M and ϕ is bounded on the closed interval $[-M, M]$.

The observation that *ϕ continuous on $(-\infty, \infty)$ and f continuous and ϕ-continuous on I do not imply f uniformly ϕ-continuous on I or f ϕ-bounded on I* is justified by the following example. Let $0 < l < \infty$ and $g(x) = lx$ for $x \in I$. Let $[a_n, b_n]$ $(a_n < b_n;\ n = 1, 2, 3, \cdots)$ be an infinite sequence of disjoint closed subintervals of I with the properties $b_1 = 1$,

$$\lim_{n\to\infty} b_n = 0; \qquad b_{n+1} < a_n, \qquad m_n + 1 \leqq m_{n+1} \qquad (n = 1, 2, 3, \cdots),$$

where

$$m_n = [lb_n - (-la_n)]/(b_n - a_n) = l(b_n + a_n)/(b_n - a_n).$$

Let

$$f(x) = \begin{cases} 0 & \text{for } x = 0, \\ lb_n & \text{for } x = b_n, \\ -la_n & \text{for } x = a_n, \end{cases} \qquad (n = 1, 2, 3, \cdots),$$

and on each closed interval $[a_n, b_n]$, $[b_{n+1}, a_n]$ $(n = 1, 2, 3, \cdots)$ let f be linear. Let $\phi \in C((-\infty, \infty))$ be arbitrary save for the satisfaction of the conditions

$$\phi(m_n) = 1/(b_n - a_n) \qquad (n = 1, 2, 3, \cdots).$$

Then f is easily seen to be continuous and ϕ-continuous on I but not uniformly ϕ-continuous on I. If p_n $(n = 1, 2, 3, \cdots)$ is any infinite sequence of numbers with

$$m_n < p_n < m_{n+1} \qquad (n = 1, 2, 3, \cdots),$$

it is clear that ϕ can be so defined on the points p_n (with $\lim_{n\to\infty} \phi(p_n) = \infty$) that f will not be ϕ-bounded on I.

In the subsequent discussion we shall have occasion at times to subject ϕ to one or more of the following conditions: continuity (usually on $-\infty < y < \infty$), convexity (usually on $-\infty < y < \infty$),

$$\liminf_{|y|\to\infty} \phi(y)/|y| > 0, \tag{11.6}$$

(11.7) $$\lim_{|y|\to\infty} \phi(y)/|y| = \infty,$$

the last two being suggested, respectively, by the particular cases $\phi(y)=|y|$, $\phi(y)=|y|^p$ $(p>1)$ on $-\infty<y<\infty$.

The function $\phi(y)=|y|^p$ $(p>1)$ is convex and satisfies (11.7). For this function ϕ it is well known(8) that ϕ-bounded variation of f implies ϕ-absolute continuity of f, so that the two conditions are equivalent; also that a necessary and sufficient condition for f to be the indefinite integral of a function g with $g\in L_\phi=L_p$ is that f be ϕ-absolutely continuous. It seems to us of some interest to determine (1) whether these properties are enjoyed by a more general class of functions L_ϕ, and (2) whether convexity of ϕ or the satisfaction of (11.7) by ϕ is the more important contributing factor in bringing about these properties.

One may easily prove

(11.8) LEMMA. *The condition* (11.6) *is equivalent to the condition that there exist positive numbers a, b with $a+b\phi(y)\geqq|y|$ for $-\infty<y<\infty$.*

Using this lemma one obtains at once

(11.9) THEOREM. *If ϕ satisfies* (11.6) *and f is ϕ-bounded [or ϕ-continuous, or of ϕ-bounded variation, or ϕ-absolutely continuous] on I, f is bounded [or continuous, or of bounded variation, or absolutely continuous, respectively] on I.*

More generally, in fact, if ϕ_1, ϕ_2 are such that there exist positive numbers a, b with $a+b\phi_1(y)\geqq\phi_2(y)$ for $-\infty<y<\infty$, ϕ_1-boundedness or ϕ_1-continuity or and so on of f on I implies the corresponding ϕ_2-property of f on I.

(11.10) THEOREM. *If ϕ satisfies* (11.7) *and f is of ϕ-bounded variation on I, f is absolutely continuous on I.*

Proof. Let $T_I^\phi(f)=B$; let $\delta_1>0$ be such that

$$\sum_{j=1}^{N} \phi\left[\frac{f(x_j)-f(x_{j-1})}{x_j-x_{j-1}}\right](x_j-x_{j-1}) < 2B \qquad \text{for } \delta<\delta_1;$$

let $M=M(\epsilon)>0$ be such that $3B/M<\epsilon$; and let $Q=Q(M)>0$ be such that

$$\phi(y)/|y| > M \qquad \text{for } |y|\geqq Q.$$

Then we have

$$MQ+\phi(y) > M|y| \qquad \text{for } -\infty<y<\infty.$$

If the norm δ of the intervals $[a_j, b_j]$ is $<\delta_1$, we have

(8) See, for example, Titchmarsh, *The theory of functions*, Oxford, 1932, pp. 384–386.

$$2B > \sum_j \phi\left[\frac{f(b_j) - f(a_j)}{b_j - a_j}\right](b_j - a_j)$$
$$> \sum_j [M \mid f(b_j) - f(a_j) \mid - MQ(b_j - a_j)]$$
$$= M \sum_j \mid f(b_j) - f(a_j) \mid - MQ \sum_j (b_j - a_j);$$

and if $\sum_j(b_j - a_j)$ is $< B/(MQ)$, we obtain

$$M \sum_j \mid f(b_j) - f(a_j) \mid < 2B + MQB/(MQ) = 3B,$$

and

$$\sum_j \mid f(b_j) - f(a_j) \mid < 3B/M < \epsilon.$$

REMARK. If ϕ is convex on $(-\infty, \infty)$ and satisfies (11.7), and f is absolutely continuous and ϕ-continuous on I, it does not follow that f is uniformly ϕ-continuous or of ϕ-bounded variation on I. This is shown by the example $\phi(y) = y^2$ on $(-\infty, \infty)$,

$$f(x) = \begin{cases} 1/2^n & \text{for } x = 1/2^n, \\ 0 & \text{for } x = 1/2^n \pm 1/2^{2n}, \qquad (n = 3, 4, 5, \cdots); \\ 0 & \text{for } x = 0, 1; \\ \text{linear on each closed interval between consecutive points for} \\ \quad \text{which } f \text{ is already defined.} \end{cases}$$

(11.11) THEOREM. *If ϕ is continuous on the closure of the set of numbers (11.1) and f' exists (finite) almost everywhere in I, we have $\int_I \phi : f' \leqq T_I^\phi(f) \leqq \infty$.*

Proof. Let S_n $(n = 1, 2, 3, \cdots)$ represent a sequence of points of subdivision of I,

$$S_n: \quad 0 = x_{n,0} < x_{n,1} < \cdots < x_{n,N_n} = 1,$$

with

$$\lim_{n\to\infty} \delta_n = 0, \qquad \delta_n = \max_{1 \leqq j \leqq N_n} (x_{n,j} - x_{n,j-1}),$$

$$T_I^\phi(f) = \lim_{n\to\infty} \sum_{j=1}^{N_n} \phi\left[\frac{f(x_{n,j}) - f(x_{n,j-1})}{x_{n,j} - x_{n,j-1}}\right](x_{n,j} - x_{n,j-1}).$$

For each n let p_n stand for the polygonal function inscribed in f with $p_n(x) = f(x)$ for $x \in S_n$, so that we have

$$T_I^\phi(f) = \lim_{n\to\infty} \int_I \phi : p_n'.$$

Since f' exists almost everywhere, p_n' tends to f' almost everywhere([9]) in I. The function ϕ being continuous and non-negative on the closure of the set of numbers (11.1), a well known theorem of Fatou yields

$$\int_I \phi:f' \leqq \lim_{n\to\infty} \int_I \phi:p_n' \leqq \infty,$$

the desired conclusion.

Remark. The hypothesis of continuity on ϕ cannot be deleted from this theorem. For the example ϕ non-measurable on I, $f(x)=x^2/2$ for $x\in I$, shows that $\int_I \phi:f'$ may not then exist. And the example $\phi(y)=1$ for $-\infty<y\leqq 0$, $\phi(y)=0$ for $0<y<\infty$, f a non-decreasing singular function constant on no subinterval of I, yields $\int_I \phi:f'=1>0=T_I^\phi(f)$.

(11.12) **Theorem.** *If f is absolutely continuous on I and $\phi\in C(\mathfrak{R}(f'))$, the condition $S^*(f', \phi, I)<\infty$ implies that f is ϕ-absolutely continuous on I and $\int_I \phi:f'=T_I^\phi(f)<\infty$.*

Proof. The hypotheses imply the existence of δ satisfying the conditions

$$0<\delta\leqq 1, \qquad \sup_{F\in\Gamma_\delta(I)} \sum_{B\in F} \phi[\mathfrak{M}_B f']\,|B|<\infty.$$

For $b_j-a_j<\delta$ $(j=1, 2, 3, \cdots)$ we have

$$\begin{aligned} 0 &\leqq \sum_j \phi\left[\frac{f(b_j)-f(a_j)}{b_j-a_j}\right](b_j-a_j) = \sum_j \phi\left[\int_{a_j}^{b_j} f'/(b_j-a_j)\right](b_j-a_j) \\ &\leqq \sup_{F\in\Gamma_\delta(H)} \sum_{B\in F} \phi[\mathfrak{M}_B f']\,|B|, \end{aligned}$$

where

$$H=\sum_j [a_j, b_j].$$

By Theorem (4.4), this sup tends to zero with $|H|$; hence the first conclusion follows. The hypotheses also imply, by Theorem (5.11),

$$\begin{aligned} \int_I \phi:f' &= \lim_{\delta\to 0} \sum_{j=1}^N \phi\left[\int_{x_{j-1}}^{x_j} f'/(x_j-x_{j-1})\right](x_j-x_{j-1}) \\ &= \lim_{\delta\to 0} \sum_{j=1}^N \phi\left[\frac{f(x_j)-f(x_{j-1})}{x_j-x_{j-1}}\right](x_j-x_{j-1}) = T_I^\phi(f)<\infty. \end{aligned}$$

From Theorem (5.13) we obtain the

Corollary. *If f is absolutely continuous on I and $\phi\in C(\mathfrak{R}(f'))$ is dominated by a convex function ψ with $\int_I \psi:f'<\infty$, the same conclusions may be drawn.*

([9]) See, for example, Titchmarsh, loc. cit. pp. 385–386.

Another immediate consequence of Theorem (5.11) is

(11.13) THEOREM. *Under the hypotheses of Theorem* (11.12) *or its Corollary, if* k *is* $\geqq 1$ *and the intervals* $[a_j, b_j]$ *satisfy the conditions*

$$a_j, b_j \in [x_{j-1}, x_j], \qquad 0 < x_j - x_{j-1} \leqq k(b_j - a_j) \qquad (j = 1, 2, \cdots, N),$$

we have

$$T_I^\phi(f) = \lim_{\delta \to 0} \sum_{j=1}^{N} \phi\left[\frac{f(b_j) - f(a_j)}{b_j - a_j}\right](x_j - x_{j-1}). \tag{11.14}$$

The next theorem is an application of Theorem (10.9).

(11.15) THEOREM. *If* f *is absolutely continuous on* I *with* $f' \in R^*(I)$, $\phi \in BC(\mathfrak{R}(f'))$, *and the intervals* $[a_j, b_j]$ *satisfy the conditions*

$$a_j, b_j \in [x_{j-1}, x_j], \quad 0 < b_j - a_j \qquad (j = 1, 2, \cdots, N),$$

we have (11.14).

The next three theorems are immediate consequences of Theorems (11.9), (11.10), and (11.11).

(11.16) THEOREM. *If* ϕ *is continuous on* $(-\infty, \infty)$ *and satisfies* (11.6), *a sufficient condition for* $f' \in L_\phi(I)$ *is that* f *be of* ϕ*-bounded variation on* I.

Proof. By Theorem (11.9), f is of bounded variation on I and f' consequently exists (finite) almost everywhere in I; since f is of ϕ-bounded variation, Theorem (11.11) yields $\int_I \phi:f' \leqq T_I^\phi(f) < \infty$.

Similarly we obtain

(11.17) THEOREM. *If* ϕ *is continuous on* $(-\infty, \infty)$ *and satisfies* (11.6), *a sufficient condition for*

$$f' \in L_\phi(I), \qquad f(x) = f(0) + \int_0^x f'(t)dt \quad \text{for} \quad x \in I, \tag{11.18}$$

is that f *be* ϕ*-absolutely continuous on* I.

(11.19) THEOREM. *If* ϕ *is continuous on* $(-\infty, \infty)$ *and satisfies* (11.7), *a sufficient condition for* (11.18) *is that* f *be of* ϕ*-bounded variation on* I.

Proof. By Theorem (11.10), f is absolutely continuous on I; and from Theorem (11.11) we have $\int_I \phi:f' \leqq T_I^\phi(f) < \infty$.

That no one of the sufficient conditions given in Theorems (11.16), (11.17), and (11.19) is necessary is shown by Example B of §8. For, let the function called f in that example be designated now by f' and let f be defined on I by

$$f(x) = \int_0^x f'(t)dt \qquad \text{for } x \in I.$$

Defining $\phi(y)=16+y^2$ for $-\infty<y<0$, we obtain a non-negative ϕ which is continuous everywhere and satisfies (11.7). And although f is absolutely continuous on I, with $f'\in L_\phi(I)$, it is not of ϕ-bounded variation on I in virtue of inequality (8.1).

The following example achieves the same result in a different way, exhibiting a non-negative ϕ which is continuous everywhere and satisfies (11.7), and an f which is absolutely continuous on I, with $f'\in L_\phi(I)$, but has an "infinite ϕ-discontinuity" at $x=0$.

Let $g(x)=x^{2/3}$ for $x\in I$, so that $g'\in L_2(I)$. Let $0\leqq a<b\leqq 1$; and let $h(x)$ for $x\in[a, b]$ be linear with $h(a)=g(a)$, $h(b)=g(b)$. Simple computations then yield

$$(11.20)\qquad \begin{aligned} \int_a^b [h']^2 &= \left[1+\frac{(ab)^{1/3}}{a^{2/3}+(ab)^{1/3}+b^{2/3}}\right](b^{1/3}-a^{1/3}) < 2(b^{1/3}-a^{1/3}) \\ &< (8/3)(b^{1/3}-a^{1/3}) = 2\int_a^b [g']^2. \end{aligned}$$

In view of the fact that g is concave on I with g' increasing continuously and monotonically toward ∞ as x decreases from 1 toward 0, there exists a sequence of numbers a_n $(n=1, 2, 3, \cdots)$ decreasing monotonically toward 0 and satisfying the following conditions:

$$(11.21)\qquad \begin{aligned} &a_1 = 1, \\ &0 < a_2 < a_1, \quad l_1 = \frac{g(a_1)-g(a_2)}{a_1-a_2} < \frac{g(a_1)}{a_1} = m_1 = 1, \\ &0 < a_{n+1} < a_n, \quad m_{n-1} < l_n = \frac{g(a_n)-g(a_{n+1})}{a_n-a_{n+1}} < \frac{g(a_n)}{a_n} = m_n \\ &\qquad\qquad (n = 2, 3, 4, \cdots). \end{aligned}$$

Each of the sequences $\{l_n\}$, $\{m_n\}$ then increases monotonically toward ∞ with n and between each pair of consecutive numbers from either sequence occurs one and only one number from the other sequence. Hence there exists a sequence of closed intervals $\{I_n\}$ with the following property: for each n, I_n has m_n as its midpoint and I_n contains neither l_n nor l_{n+1}. That is, the set $\sum_{n=1}^{\infty} I_n$ contains no point of the sequence $\{l_n\}$.

Let f be defined on I as follows: $f(0)=0$; for each n, $f(x)$ is linear on the interval $[a_{n+1}, a_n]$ with $f(a_{n+1})=g(a_{n+1})$, $f(a_n)=g(a_n)$. From inequality (11.20) we infer $f'\in L_2(I)$. Since f is monotone and continuous on I, and on each interval $[\epsilon, 1]$ $(0<\epsilon<1)$ it is polygonal and so absolutely continuous, it is absolutely continuous on I.

For y not in $\sum_{n=1}^{\infty} I_n$, let $\phi(y)=y^2$; for $y=m_n$ $(n=1, 2, 3, \cdots)$, let $\phi(y)=y^4$; in each closed half (that is, left and right half) of each interval I_n, let ϕ be

linear. Then ϕ is continuous for $-\infty < y < \infty$ and satisfies (11.7). We also have, by aid of (11.20),

$$\int_I \phi:f' = \sum_{n=1}^{\infty} \int_{a_{n+1}}^{a_n} \phi:f' = \sum_{n=1}^{\infty} \int_{a_{n+1}}^{a_n} [f']^2 < 2 \sum_{n=1}^{\infty} \int_{a_{n+1}}^{a_n} [g']^2 = 2 \int_I [g']^2 = 8/3.$$

On the other hand if x tends to 0 over the sequence $\{a_n\}$, we have

$$f(a_n)/a_n = a_n^{2/3}/a_n = 1/a_n^{1/3} = m_n, \qquad \phi(m_n) = m_n^4 = 1/a_n^{4/3},$$

$$\phi\left[\frac{f(a_n) - f(0)}{a_n - 0}\right](a_n - 0) = \phi(m_n)a_n = 1/a_n^{1/3} \to \infty.$$

From the Corollary to Theorem (5.13) we obtain at once

(11.22) THEOREM. *If f is absolutely continuous on I and ϕ is convex and continuous on $\Re(f')$, we have $\int_I \phi:f' = T_I^\phi(f) \leqq \infty$.*

Using Theorem (11.12), we obtain the

COROLLARY. *If f is absolutely continuous on I, ϕ is convex and continuous on $\Re(f')$, and f is of ϕ-bounded variation on I, then f is ϕ-absolutely continuous on I.*

Theorem (11.22) implies that if f is absolutely continuous on I and ϕ is convex and continuous on $\Re(f')$, we have

$$T_I^\phi(f) = \lim_{\epsilon\to 0} T_{[\epsilon,1]}^\phi(f) \leqq \infty \qquad (0 < \epsilon < 1). \tag{11.23}$$

That the assumption of convexity on ϕ is essential for (11.23) may be seen from the following example, which exhibits an f which is absolutely continuous on I and a ϕ which is continuous everywhere and satisfies (11.7), but f has a "finite ϕ-discontinuity" at $x=0$, with $\limsup_{x\to 0} \phi[(f(x)-f(0))/x]x = 1$; and for each positive $\epsilon < 1$

$$f \text{ is } \phi\text{-absolutely continuous on } [\epsilon, 1],$$

$$T_{[\epsilon,1]}^\phi(f) < 4\cdot 2^{1/2}, \qquad T_I^\phi(f) = T_{[0,\epsilon]}^\phi(f) = \infty.$$

Let $g(x) = x^{1/2}$ for $x \in I$, so that $g' \in L_{3/2}(I)$. Let $0 \leqq a < b \leqq 1$; and let $h(x)$ for $x \in [a, b]$ be linear with $h(a) = g(a)$, $h(b) = g(b)$. Simple computations provide the inequality

$$\begin{aligned} \int_a^b [h']^{3/2} &= \frac{b^{1/4} + a^{1/4}}{(b^{1/2} + a^{1/2})^{1/2}} (b^{1/4} - a^{1/4}) < 2(b^{1/4} - a^{1/4}) \\ &= 2^{1/2} \int_a^b [g']^{3/2}. \end{aligned} \tag{11.24}$$

As in the example last described, g is concave on I with g' increasing continuously and monotonically toward ∞ as x decreases from 1 toward 0, and there exists a sequence of numbers $\{a_n\}$ satisfying conditions (11.21). This time let I_n $(n=1, 2, 3, \cdots)$ be a sequence of closed intervals with the following property: for each n, I_n has l_n as its midpoint and I_n contains neither m_{n-1} nor m_n.

Let f be defined on I as in the preceding example, so that it is absolutely continuous there. From (11.24) it is clear that $f'\in L_{3/2}(I)$. For y not in $\sum_{n=1}^{\infty} I_n$, let $\phi(y)=y^2$; for $y=l_n$ $(n=1, 2, 3, \cdots)$, let $\phi(y)=y^{3/2}$; in each closed half (that is, left and right half) of each interval I_n, let ϕ be linear. Thus ϕ is continuous for $-\infty<y<\infty$ and satisfies (11.7).

On the interval $[\epsilon, 1]$, f satisfies a Lipschitz condition; let $M=M(\epsilon)$ satisfy the inequality

$$|f(x_1)-f(x_2)| \leqq M\,|x_1-x_2| \qquad \text{for } x_1, x_2 \in [\epsilon, 1].$$

On the interval $[0, M]$, ϕ is bounded; hence ϕ is dominated on the interval $(-\infty, \infty)$ by a convex function ψ, with $\int_\epsilon^1 \psi:f'<\infty$. From the corollary to Theorem (11.12) we therefore infer the ϕ-absolute continuity of f on $[\epsilon, 1]$ and, in view of (11.24) also, the relation

$$T^{\phi}_{[\epsilon,1]}(f) = \int_\epsilon^1 \phi:f' = \int_\epsilon^1 [f']^{3/2} < \int_I [f']^{3/2} < 2^{1/2}\int_I [g']^{3/2} = 4\cdot 2^{1/2}.$$

For $0<x\leqq 1$ we also have $f(x)/x\geqq 1$ and

$$\phi[f(x)/x]x \leqq [f(x)/x]^2 x \leqq [g(x)/x]^2 x = 1,$$

whence

$$\limsup_{x\to 0} \phi[f(x)/x]x \leqq 1;$$

and if x tends to 0 over the sequence $\{a_n\}$, we have

$$f(a_n)/a_n = g(a_n)/a_n = a_n^{1/2}/a_n = 1/a_n^{1/2} = m_n, \qquad \phi(m_n) = m_n^2 = 1/a_n,$$
$$\phi[f(a_n)/a_n]a_n = 1 \qquad (n = 1, 2, 3, \cdots),$$

which implies

$$\limsup_{x\to 0} \phi[f(x)/x]x = 1.$$

To show the property $T^{\phi}_{[0,\epsilon]}=\infty$ $(0<\epsilon<1)$, we choose a particular sequence $\{a_n\}$; namely $a_n=1/2^{2(n-1)}$ $(n=1, 2, 3, \cdots)$. This clearly satisfies the principal inequality of (11.21),

$$g(a_{n-1})/a_{n-1} < [g(a_n)-g(a_{n+1})]/(a_n-a_{n+1}) < g(a_n)/a_n,$$

which reduces to

$$1/a_{n-1}^{1/2} < 1/(a_n^{1/2} + a_{n+1}^{1/2}) < 1/a_n^{1/2},$$

since we now have

$$a_n^{1/2} + a_{n+1}^{1/2} = 1/2^{n-1} + 1/2^n = (1/2^{n-2})(1/2 + 1/4) < 1/2^{n-2} = a_{n-1}^{1/2}.$$

From the obvious inequality

$$[g(a_n) - g(a_{n+2})]/(a_n - a_{n+2}) < g(a_n)/a_n = m_n,$$

we infer that a line drawn from the midpoint P_{n+1} of the segment of f standing over the interval $[a_{n+2}, a_{n+1}]$, with slope m_n, meets the segment of f over $[a_{n+1}, a_n]$ at a point Q_n to the left of the midpoint P_n of that segment. Thus if c_n, d_n are the abscissae of P_{n+1}, Q_n, respectively, it is clear that the intervals $[c_n, d_n]$ $(n=1, 2, 3, \cdots)$ are disjoint, with

$$(a_n - a_{n+2})/2 > d_n - c_n > a_{n+1} - c_n = (a_{n+1} - a_{n+2})/2.$$

But for each n we have

$$[f(d_n) - f(c_n)]/(d_n - c_n) = m_n = g(a_n)/a_n = 1/a_n^{1/2} = 2^{n-1},$$

$$\phi(m_n) = m_n^2 = 2^{2(n-1)},$$

$$\phi(m_n)(d_n - c_n) > \phi(m_n)(a_{n+1} - a_{n+2})/2 = 2^{2(n-1)}(1/2^{2n} - 1/2^{2(n+1)})/2 = 3/32.$$

Since $d_n - c_n$ is $<(a_n - a_{n+2})/2$, which tends to 0 with $1/n$, and the intervals $[c_n, d_n]$ $(n=1, 2, 3, \cdots)$ are disjoint, it follows that in any interval $[0, \epsilon]$, and for an arbitrary norm $\delta>0$, the sum whose lim sup is $T^{\phi}_{[0,\epsilon]}(f)$ can be made arbitrarily large by a suitable choice of points of subdivision of $[0, \epsilon]$.

We conclude with two more theorems which show that well known properties of the L_p $(p \geqq 1)$ classes of functions hold for certain more general L_ϕ classes.

(11.25) THEOREM. *If ϕ is convex on $(-\infty, \infty)$ and satisfies* (11.6), *a necessary and sufficient condition for* (11.18) *is that f be ϕ-absolutely continuous on I.*

The sufficiency of the condition follows from Theorem (11.17), the necessity from the Corollary to Theorem (11.12).

(11.26) THEOREM. *If ϕ is convex on $(-\infty, \infty)$ and satisfies* (11.7), *a necessary and sufficient condition for* (11.18) *is that f be of ϕ-bounded variation on I.*

The sufficiency is a consequence of Theorem (11.19); the necessity follows from Theorem (11.22).

BROWN UNIVERSITY,
PROVIDENCE, R. I.
THE UNIVERSITY OF CALIFORNIA,
BERKELEY, CALIF.

ON SOME SINGULAR MONOTONIC FUNCTIONS WHICH ARE STRICTLY INCREASING

BY

R. SALEM

1. A continuous non-decreasing function $f(x)$ defined for $0 \leqq x \leqq 1$ ($f(0)=0$, $f(1)=1$) and which is purely singular, that is to say, which has the property

$$df/dx = 0$$

almost everywhere, may be constant in every interval contiguous to a perfect set of measure zero: it is usually said, in this case, that $f(x)$ is of the Cantor type. There are, however, monotonic continuous functions, purely singular, which are increasing in the strict sense, that is, $f(x') > f(x)$ whenever $x' > x$.

While the existence of functions of the Cantor type is almost intuitive and their construction is immediate by successive approximations, the existence of strictly increasing singular functions lies deeper. Actually, if we except Minkowski's function $?(x)$, of which we shall speak later (and whose singularity is by no means obvious), no simple direct construction of such functions seems to be known. Functions of this type usually have been obtained by "convolutions" of functions of the Cantor type and the proof that they are singular strictly increasing functions is somewhat difficult(1). Thus, it seems to be of interest to give simple direct constructions of strictly increasing singular functions.

2. Let us consider, in the plane, the straight line PQ joining the point P of cartesian coordinates x, y, to the point Q of cartesian coordinates $x+\Delta x$, $y+\Delta y$, $\Delta x>0$, $\Delta y>0$. Let λ_0, λ_1 be two numbers, essentially positive, such that $\lambda_0+\lambda_1=1$ ($\lambda_0 \neq \lambda_1$). Let us now consider the point R whose coordinates are

$$x + \Delta x/2, \qquad y + \lambda_0 \Delta y,$$

that is to say, the horizontal distance between P, R or between Q, R is $\Delta x/2$, while the vertical distance between P, R is $\lambda_0 \Delta y$, and the vertical distance between R, Q is $\lambda_1 \Delta y$. If we replace the straight line PQ by the broken line PRQ, we will say that we perform on PQ the transformation $T(\lambda_0, \lambda_1)$.

Presented to the Society, November 28, 1942; received by the editors October 8, 1942.

(1) See for example Jessen and Wintner, *Distribution functions and the Riemann zeta function*, Trans. Amer. Math. Soc. vol. 38 (1935) pp. 48–88 and particularly p. 61; Kershner and Wintner, *On symmetric Bernoulli convolutions*, Amer. J. Math. vol. 57 (1935) pp. 541–548; Wiener and Wintner, *Fourier-Stieltjes transforms and singular infinite convolutions*, Amer. J. Math. vol. 60 (1938) pp. 513–522 and particularly p. 521. For earlier examples see Denjoy, J. Math. Pures Appl. 1915 pp. 204–209 (which was the first example given); Sierpinski, Giornale di Matematiche vol. 54 (1916) pp. 314–334; Rajchman, Fund. Math. vol. 2 (1921) pp. 50–63.

Definition of a function $f(x)$. Let now $f_0(x)$ be for $0 \leqq x \leqq 1$ the function equal to x, that is to say represented by the straight line OA joining the origin O to the point $A(1, 1)$. Let us perform on OA the transformation $T(\lambda_0, \lambda_1)$. We get a broken line consisting of two straight lines and representing an increasing function $f_1(x)$. Let us perform on each of those two straight lines the transformation $T(\lambda_0, \lambda_1)$. We get a broken line consisting of 2^2 straight lines and representing an increasing function $f_2(x)$. Proceeding in the same way we get after p operations a function $f_p(x)$ strictly increasing ($f_p(0) = 0, f_p(1) = 1$) represented by a polygonal line consisting of 2^p straight lines, the vertices having for abscissae the points $k/2^p$ ($k = 1, 2, \cdots, 2^p - 1$).

Putting

$$\max(\lambda_0, \lambda_1) = \mu \tag{1}$$

we have essentially, by our hypothesis, $\mu < 1$, and it is immediately seen that

$$|f_{p+1} - f_p| \leqq \mu^p.$$

Thus $f_p(x)$ converges uniformly to a continuous function $f(x)$ ($f(0) = 0$, $f(1) = 1$). This function $f(x)$ is strictly increasing because for every p the vertices of the curve $y = f_p(x)$ belong to the curve $y = f(x)$: thus if $f(x)$ was constant in some interval, there would be a p for which two different vertices of $y = f_p(x)$ would have the same ordinate, which is impossible.

The ordinate of the vertex of $y = f_p(x)$ whose abscissa is given by

$$\theta_1/2 + \theta_2/2^2 + \cdots + \theta_p/2^p \qquad (\theta_i = 0 \text{ or } 1)$$

is given by

$$\lambda_0[\theta_1 + \lambda_{\theta_1}\theta_2 + \lambda_{\theta_1}\lambda_{\theta_2}\theta_3 + \cdots + \lambda_{\theta_1}\lambda_{\theta_2}\cdots\lambda_{\theta_{p-1}}\theta_p],$$

and thus, by continuity, if

$$x = \theta_1/2 + \theta_2/2^2 + \cdots + \theta_p/2^p + \cdots, \tag{2}$$

we have

$$f(x) = \lambda_0[\theta_1 + \lambda_{\theta_1}\theta_2 + \lambda_{\theta_1}\lambda_{\theta_2}\theta_3 + \cdots + \lambda_{\theta_1}\lambda_{\theta_2}\cdots\lambda_{\theta_{p-1}}\theta_p + \cdots] \tag{3}$$

the series being obviously convergent. If x has two different dyadic developments, the formula (3) gives for $f(x)$ the same value.

Let us remark also that if x and $x' > x$ have the first p digits of their dyadic developments identical, and equal to $\theta_1, \theta_2, \cdots, \theta_p$, then

$$f(x') - f(x) < \lambda_{\theta_1}\lambda_{\theta_2}\cdots\lambda_{\theta_p}. \tag{4}$$

This is seen immediately by the formula (3) or geometrically.

Proof that $f(x)$ is singular. We shall now prove that the function $f(x)$ is singular.

It is well known that almost all numbers in (0, 1) are "normal" in the scale of 2, that is, are such that

$$\theta_1 + \theta_2 + \cdots + \theta_p = p/2 + o(p) \qquad \text{when} \quad p \to \infty.$$

Let N be the set of these normal numbers. We have meas $N=1$. Let us fix an x belonging to N. Let x be given by (2). p being a positive integer the number $x+\epsilon_{p+1}/2^{p+1}$, where

$$\epsilon_{p+1} = 1 \quad \text{if} \quad \theta_{p+1} = 0,$$
$$\epsilon_{p+1} = -1 \quad \text{if} \quad \theta_{p+1} = 1,$$

has a dyadic development whose first p digits are the same as for x, that is, $\theta_1, \theta_2, \cdots, \theta_p$.

Hence by (4)

$$|f(x + \epsilon_{p+1}/2^{p+1}) - f(x)| < \lambda_{\theta_1}\lambda_{\theta_2} \cdots \lambda_{\theta_p}.$$

Now x being normal

$$\theta_1 + \theta_2 + \cdots + \theta_p = p/2 + \phi(p)$$

with $|\phi(p)|/p \to 0$ when $p \to \infty$. Hence

$$\lambda_{\theta_1}\lambda_{\theta_2} \cdots \lambda_{\theta_p} = \lambda_0^{p/2-\phi(p)}\lambda_1^{p/2+\phi(p)} < (\lambda_0\lambda_1)^{p/2-|\phi(p)|},$$

hence

$$2^{p+1}|f(x + \epsilon_{p+1}/2^{p+1}) - f(x)| < (2(\lambda_0\lambda_1)^{1/2})^p \cdot 2/(\lambda_0\lambda_1)^{|\phi(p)|}. \tag{5}$$

Now λ_0 and λ_1 being essentially different and $\lambda_0+\lambda_1$ being equal to 1, we have

$$2(\lambda_0\lambda_1)^{1/2} < 1.$$

This, together with lim $|\phi(p)|/p=0$ proves that the second member of (5) tends to zero for $p=\infty$, and thus, *if $f(x)$ has a derivative at the point x*, this derivative cannot have a value different from zero. But by a classical theorem $f'(x)$ exists and is finite almost everywhere, hence almost everywhere in N. Hence, also almost everywhere, $f'(x)=0$ which proves our theorem.

Modulus of continuity of $f(x)$. The vertical distance between two vertices of abscissae $k/2^p$, $k+1/2^p$ being less than μ^p where μ is defined by (1), we have immediately that if $1/2^{p+1} \leqq x'-x < 1/2^p$ then

$$f(x') - f(x) < 2\mu^p \leqq 2\mu^{-1}\mu^{\log (x'-x)^{-1}/\log 2} = 2\mu^{-1}(x'-x)^{|\log \mu|/\log 2}.$$

Hence $f(x)$ satisfies a Lipschitz condition of order $|\log \mu|/\log 2$.

Fourier-Stieltjes coefficients of $f(x/2\pi)$. Let

$$c_n = \int_0^{2\pi} e^{nix} df(x/2\pi).$$

We divide the y-axis by the points of subdivision corresponding to the vertices existing at the pth stage of the construction of the function and we observe that the vertical distance between the vertex whose abscissa is

$$2\pi(\theta_1/2 + \theta_2/2^2 + \cdots + \theta_p/2^p)$$

and the following one, which can be written

$$2\pi\left(\theta_1/2 + \theta_2/2^2 + \cdots + \theta_p/2^p + \sum_{s=p+1}^{\infty} (1/2^s)\right),$$

has the value $\lambda_{\theta_1}\lambda_{\theta_2} \cdots \lambda_{\theta_p}$. We thus get for approximate expression of the integral

$$\sum \lambda_{\theta_1}\lambda_{\theta_2} \cdots \lambda_{\theta_p} e^{2\pi n i(\theta_1/2+\theta_2/2^2+\cdots+\theta_p/2^p)},$$

the summation being extended to the 2^p combinations of the values 0 and 1 of the θ_i. This sum is equal to

$$\prod_{k=1}^{p} [\lambda_0 + \lambda_1 e^{2\pi n i/2^k}]$$

and thus, making $p = \infty$, we have

$$c_n = \prod_{k=1}^{\infty} [\lambda_0 + \lambda_1 e^{2\pi n i/2^k}].$$

We can also write

$$c_n = \prod_{k=1}^{\infty} e^{\pi n i/2^k}[\lambda_0 e^{-\pi n i/2^k} + \lambda_1 e^{\pi n i/2^k}] = e^{\pi n i}\prod_{k=1}^{\infty} [\lambda_0 e^{-\pi n i/2^k} + \lambda_1 e^{\pi n i/2^k}],$$

or, putting $\lambda_0 = (1-r)/2$, $\lambda_1 = (1+r)/2$

$$c_n = e^{\pi n i}\prod_{k=1}^{\infty} [\cos(\pi n/2^k) + ir\sin(\pi n/2^k)]$$

that gives

$$|c_n|^2 = \prod_{k=1}^{\infty} [\cos^2(\pi n/2^k) + r^2 \sin^2(\pi n/2^k)].$$

If we take $n = 2^m$, we have

$$|c_{2^m}|^2 > r^2 \cos^2(\pi/4)\cos^2(\pi/8)\cos^2(\pi/16)\cdots$$

and thus c_n does not tend to zero for $n = \infty$.

3. **Generalization of the preceding function.** Instead of constructing our function with an infinity of identical transformations $T(\lambda_0, \lambda_1)$, let us change the transformation used at every step of the construction.

Thus $F_0(x)$ being equal to x in $(0, 1)$ let O be the point $(0, 0)$, A the point

(1, 1) and let us perform on OA the transformation $T(\lambda_0^{(1)}, \lambda_1^{(1)})$. We get a broken line consisting of two straight lines and representing $F_1(x)$. On each of those two straight lines we perform the transformation $T(\lambda_0^{(2)}, \lambda_1^{(2)})$; on the 2^2 straight lines constituting $F_2(x)$ we perform the transformation $T(\lambda_0^{(3)}, \lambda_1^{(3)})$ to get $F_3(x)$, and so on.

Let

$$\lambda_0^{(k)} = (1 - r_k)/2, \quad \lambda_1^{(k)} = (1 + r_k)/2.$$

We assume that $-1 < r_k < 1$ for every k and that if we put

$$u_p = \prod_{k=1}^{p} \left(\frac{1 + |r_k|}{2} \right)$$

the series $\sum u_p$ converges. (This is certainly the case, for example, if $-\alpha < r_k < \alpha$, $0 < \alpha < 1$, but can be secured under less stringent conditions.) Then there is no change in the argument used in §2 to prove that $F_p(x)$ tends uniformly to a continuous function $F(x)$ strictly increasing from 0 to 1 in the interval (0, 1).

In the same way as before we prove that if

$$x = \theta_1/2 + \theta_2/2^2 + \cdots + \theta_p/2^p + \cdots, \tag{6}$$

we have

$$F(x) = \theta_1 \lambda_0^{(1)} + \theta_2 \lambda_{\theta_1}^{(1)} \lambda_0^{(2)} + \theta_3 \lambda_{\theta_1}^{(1)} \lambda_{\theta_2}^{(2)} \lambda_0^{(3)} + \cdots. \tag{7}$$

Finally, if x and $x' > x$ have the same first p digits in their dyadic development, we have

$$F(x') - F(x) < \lambda_{\theta_1}^{(1)} \lambda_{\theta_2}^{(2)} \cdots \lambda_{\theta_p}^{(p)} = (1/2^p) \prod_{k=1}^{p} (1 - \epsilon_k r_k) \tag{8}$$

where

$$\epsilon_k = 1 \quad \text{if} \quad \theta_k = 0, \qquad \epsilon_k = -1 \quad \text{if} \quad \theta_k = 1.$$

It will be useful to observe that if x is given by (6) we can also write

$$\lambda_{\theta_k}^{(k)} = (1/2)(1 - \epsilon_k r_k) = (1/2)(1 - \phi_k(x) r_k)$$

where $\{\phi_k(x)\}$ denotes the system of Rademacher's functions ($k = 1, 2, \cdots$). Thus, with this notation, the inequality (8) is written

$$F(x') - F(x) < (1/2^p) \prod_{k=1}^{p} (1 - \phi_k(x) r_k).$$

We can now prove the following theorem.

THEOREM. *The function $F(x)$ is purely singular when, and only when, the series $\sum r_k^2$ diverges.*

We shall make use of the following theorem, due to Zygmund([2]): for almost all x we have

$$\liminf \sum_{1}^{n} - r_k\phi_k(x) = -\infty$$

if the series $\sum r_k^2$ diverges.

We deduce immediately from this result and from the inequality

$$1 - r_k\phi_k(x) < e^{-r_k\phi_k(x)}$$

that for almost all x

$$\liminf_{p=\infty} \prod_{k=1}^{p} (1 - r_k\phi_k(x)) = 0 \tag{9}$$

provided that $\sum r_k^2 = \infty$.

The proof of the first part of our theorem is now immediate. Taking an x belonging to the set E (meas $E=1$) for which (9) holds, we have

$$\left| F(x + (\epsilon_{p+1}/2^{p+1})) - F(x) \right| < (1/2^p) \prod_{k=1}^{p} (1 - \phi_k(x)r_k),$$

hence

$$\liminf 2^{p+1} \left| F(x + (\epsilon_{p+1}/2^{p+1})) - F(x) \right| = 0$$

and if $F'(x)$ exists it is equal to zero. The proof is completed as above.

To prove the second part of our theorem let us suppose that $\sum r_k^2 < \infty$. We know by a classical theorem that in this case the series $\sum r_k\phi_k(x)$ converges in a set $\mathcal{E}$ of measure 1. From this and from the hypothesis $\sum r_k^2 < \infty$ it is easy to deduce that the infinite product

$$\prod_{1}^{\infty} (1 - r_k\phi_k(x)) \tag{10}$$

is convergent when x belongs to $\mathcal{E}$. Fixing an x belonging to $\mathcal{E}$, let us remark that ϵ_{p+1} having the same signification as above the dyadic developments of x and $x+\epsilon_{p+1}/2^{p+1}$ have all their digits equal except the digits of rank $p+1$. From this it is easy to deduce, for example geometrically, that

$$\frac{1 - |r_{p+1}|}{2} \lambda_{\theta_1}^{(1)} \cdots \lambda_{\theta_p}^{(p)} < \left| F\left(x + \frac{\epsilon_{p+1}}{2^{p+1}}\right) - F(x) \right| < \frac{1 + |r_{p+1}|}{2} \lambda_{\theta_1}^{(1)} \cdots \lambda_{\theta_p}^{(p)}.$$

([2]) Zygmund, *On lacunary trigonometric series*, Trans. Amer. Math. Soc. vol. 34 (1932) p. 435. The proof given there for lacunary trigonometric series is immediately applicable to Rademacher's functions.

Now the first part of this inequality together with the convergence of the product (10) shows that

$$\liminf 2^{p+1}\left|F\left(x+\frac{\epsilon_{p+1}}{2^{p+1}}\right)-F(x)\right|>0.$$

Hence, whenever $F'(x)$ exists for $x\in\mathcal{E}$, $F'(x)$ is not zero. Remembering that $F'(x)$ exists and is finite almost everywhere, we have that $F'(x)\neq 0$ almost everywhere, and thus $F(x)$ cannot be purely singular. This completes the proof of the theorem.

Modulus of continuity of $F(x)$. The argument is the same as before. If

$$1/2^{p+1}\leqq x'-x<1/2^p,$$

we have

$$F(x')-F(x)<2\prod_{k=1}^{p}\left(\frac{1+|r_k|}{2}\right).$$

Thus if $\omega(\delta)$ is the modulus of continuity, we have

$$\omega(\delta)<4\prod_{k=1}^{(|\log\delta|/\log 2)}\left(\frac{1+|r_k|}{2}\right).$$

Fourier-Stieltjes coefficients of $F(x/2\pi)$. There is no change in the argument used above for $f(x)$ to prove that if

$$c_n=\int_0^{2\pi}e^{nix}dF\left(\frac{x}{2\pi}\right)$$

we have for approximate expression of c_n

$$\sum\lambda_{\theta_1}^{(1)}\cdots\lambda_{\theta_p}^{(p)}e^{2\pi ni(\theta_1/2+\theta_2/2^2+\cdots+\theta_p/2^p)}$$

the summation being extended to the 2^p combinations of the values 0 and 1 of the θ_i. We get thus

$$\begin{aligned}c_n&=\prod_{k=1}^{\infty}\left[\lambda_0^{(k)}+\lambda_1^{(k)}e^{2\pi ni/2^k}\right]\\&=e^{\pi ni}\prod_{k=1}^{\infty}\left[\lambda_0^{(k)}e^{-\pi ni/2^k}+\lambda_1^{(k)}e^{\pi ni/2^k}\right]\\&=e^{\pi ni}\prod_{k=1}^{\infty}\left[\cos(\pi n/2^k)+ir_k\sin(\pi n/2^k)\right]\end{aligned}$$

and

$$|c_n|^2=\prod_{k=1}^{\infty}\left[\cos^2(\pi n/2^k)+r_k^2\sin^2(\pi n/2^k)\right].$$

It is immediately seen—as in the case of $f(x)$—that if r_k does not tend to zero, we have $|c_n| \neq o(1)$.

4. **The Minkowski function** $?(x)$. This function was defined by Minkowski[3] for the purpose of establishing a one-one correspondence between the rational numbers of (0, 1) and the quadratic irrationals of (0, 1). The properties of the function have been recently investigated by Denjoy[4] who has proved that it is purely singular and given other important properties and generalizations of Minkowski's function.

We propose to give here some new indications about this function, concerning particularly its modulus of continuity and its Fourier-Stieltjes coefficients. For the sake of completeness we shall give the definition of the function and the proof of its singularity.

Definition of the function $?(x)$. We define first

$$?(0) = ?(0/1) = 0, \qquad ?(1) = ?(1/1) = 1.$$

We next take the "mediant" $1/2=(0+1)/(1+1)$ of the two Farey fractions 0/1 and 1/1 and we define $?(0+1/1+1)$ to be the arithmetic mean between $?(0)$ and $?(1)$, that is, 1/2.

We define in the same way

$$?(1/3) = ?\left(\frac{0+1}{1+2}\right) = \frac{?(0) + ?(1/2)}{2} = 1/4,$$

$$?(2/3) = ?\left(\frac{1+1}{2+1}\right) = \frac{?(1/2) + ?(1)}{2} = 3/4.$$

Generally if, by this process, we have defined $?(p/q)$ and $?(p'/q')$ for two consecutive irreducible fractions p/q, p'/q', we define

$$?\left(\frac{p+p'}{q+q'}\right) = \frac{?(p/q) + ?(p'/q')}{2}.$$

At the nth stage the function is defined for 2^n+1 values of x and the ordinates corresponding to these values of x are of the form $k/2^n$ $(k=0, 1, 2, \cdots, 2^n)$. The definition of $?(x)$ for every x follows by continuity.

Let now x be a rational number put in the form of a finite continued fraction:

$$x = (a_0, a_1, \cdots, a_n), \qquad a_0 = 0 \qquad\qquad (0 \leq x \leq 1).$$

Let p_0/q_0, p_1/q_1, $\cdots$, $p_n/q_n = x$ be the successive convergents ($p_0/q_0=0$, $p_1/q_1=1/a_1$, $\cdots$). Let us assume that at a certain stage (the mth) of the

(3) H. Minkowski, Gesammelte Abhandlungen vol. 2 (1911) pp. 50–51.

(4) A. Denjoy, C. R. Acad. Sci. Paris vol. 194 (1932) pp. 44–45 and J. Math. Pures Appl. vol. 17 (1938) pp. 105–151.

construction of $?(x)$ the fractions p_{k-2}/q_{k-2}, p_{k-1}/q_{k-1} are consecutive. (This happens certainly for $p_0/q_0=0$ and $p_1/q_1=1/a_1$, $1/a_1$ appearing as consecutive to $0/1$ at the (a_1-1)th stage of the construction.) Let

$$y_{k-2} = ?\left(\frac{p_{k-2}}{q_{k-2}}\right), \qquad y_{k-1} = ?\left(\frac{p_{k-1}}{q_{k-1}}\right).$$

By definition

$$?\left(\frac{p_{k-1}+p_{k-2}}{q_{k-1}+q_{k-2}}\right) = \frac{y_{k-1}+y_{k-2}}{2}.$$

Now it is well known that $(p_{k-1}+p_{k-2})/(q_{k-1}+q_{k-2})$ is irreducible and thus at the next stage (the $(m+1)$th) the fraction (irreducible) $(2p_{k-1}+p_{k-2})/(2q_{k-1}+q_{k-2})$ will appear with

$$?\left(\frac{2p_{k-1}+p_{k-2}}{2q_{k-1}+q_{k-2}}\right) = \frac{y_{k-1}+(y_{k-1}+y_{k-2})/2}{2}.$$

Continuing in the same way we see that

$$?\left(\frac{p_k}{q_k}\right) = ?\left(\frac{a_kp_{k-1}+p_{k-2}}{a_kq_{k-1}+q_{k-2}}\right)$$
$$= \frac{y_{k-1}}{2} + \frac{y_{k-1}}{2^2} + \cdots + \frac{y_{k-1}}{2^{a_k}} + \frac{y_{k-2}}{2^{a_k}},$$

hence, if we put $?(p_k/q_k)=y_k$,

$$y_k = (1-(1/2^{a_k}))y_{k-1} + y_{k-2}/2^{a_k}$$

or

$$y_k - y_{k-1} = -(1/2^{a_k})(y_{k-1}-y_{k-2}).$$

Now p_k/q_k when it appears is consecutive to p_{k-1}/q_{k-1}. Hence we can repeat the argument, and if $y_n=?(p_n/q_n)=?(x)$, we have

$$y_n - y_{n-1} = (-1/2^{a_n})(-1/2^{a_{n-1}})\cdots(-1/2^{a_2})(y_1-y_0).$$

Now $y_0=0$, $y_1=1/2^{a_1-1}$, hence

$$y_n - y_{n-1} = (-1)^{n-1}\frac{1}{2^{(a_1+\cdots+a_n)-1}}$$

and thus

$$y_n = \frac{1}{2^{a_1-1}} - \frac{1}{2^{(a_1+a_2)-1}} + \frac{1}{2^{(a_1+a_2+a_3)-1}} - \cdots + (-1)^{n-1}\frac{1}{2^{(a_1+\cdots+a_n)-1}}.$$

Now by continuity we get the following result: if

$$x = (0, a_1, a_2, \cdots, a_n, \cdots),$$

we have

$$(11)\quad ?(x) = \frac{1}{2^{a_1-1}} - \frac{1}{2^{(a_1+a_2)-1}} + \cdots + (-1)^{n-1}\frac{1}{2^{(a_1+\cdots+a_n)-1}} + \cdots,$$

and it is easy to see that if x is rational, the two different developments of x give the same $?(x)$.

From this we deduce the more elementary properties of $?(x)$, namely:

If x is rational, $?(x)$ is of the form $k/2^s$ (k, s integers).

If x is irrational, the dyadic development of $?(x)$ is infinite.

If x is a quadratic irrational $(0, a_1, a_2, \cdots)$ is periodic and thus $?(x)$, being the difference of two periodic dyadic developments, is rational.

It is not difficult to see that the reciprocals of these results are true.

The fact that $?(x)$ is strictly increasing is an immediate consequence of its construction.

Proof of the singularity of $?(x)$. Let $x=(0, a_1, \cdots, a_n, \cdots)$. We know that for almost all x $\limsup a_n=\infty$. Let N be the set of such numbers (meas. $N=1$) and let us fix an x belonging to N. Let $x=(0, a_1, \cdots, a_n, \cdots)$, let $?(x)=y$ and let

$$r_n = p_n/q_n = (0, a_1, \cdots, a_n), \qquad \rho_n = ?(r_n),$$

and let us write, as usual $a'_{n+1}=(a_{n+1}, a_{n+2}, \cdots)$. We have

$$x = \frac{a'_{n+1}p_n + p_{n-1}}{a'_{n+1}q_n + q_{n-1}}, \qquad \left| x - \frac{p_n}{q_n}\right| = \frac{1}{(a'_{n+1}q_n + q_{n-1})q_n},$$

and thus

$$\frac{1}{(a_{n+1}+2)q_n^2} < \left| x - \frac{p_n}{q_n}\right| < \frac{1}{a_{n+1}q_n^2}$$

and

$$y - \rho_n = (-1)^n\left[\frac{1}{2^{(a_1+\cdots+a_{n+1})-1}} - \frac{1}{2^{(a_1+\cdots+a_{n+2})-1}} + \cdots\right]$$

which gives

$$\frac{1}{2^{a_1+\cdots+a_{n+1}}} < | y - \rho_n | < \frac{1}{2^{(a_1+\cdots+a_{n+1})-1}}.$$

Hence, we have

$$\delta_n = \left|\frac{y-\rho_n}{x-r_n}\right| < \frac{2(a_{n+1}+2)q_n^2}{2^{a_1+\cdots+a_{n+1}}}$$

and

$$\delta_{n-1} = \left| \frac{y - \rho_{n-1}}{x - r_{n-1}} \right| > \frac{a_n q_{n-1}^2}{2^{a_1 + \cdots + a_n}}.$$

Consequently,

$$\begin{aligned} \frac{\delta_n}{\delta_{n-1}} &< 2 \frac{1}{2^{a_{n+1}}} \left(\frac{a_{n+1} + 2}{a_n} \right) \left(\frac{q_n}{q_{n-1}} \right)^2 \\ &< 2 \frac{1}{2^{a_{n+1}}} \left(\frac{a_{n+1} + 2}{a_n} \right) (a_n + 1)^2 < C \frac{a_n a_{n+1}}{2^{a_{n+1}}}, \end{aligned}$$

C being an absolute constant.

Now we can certainly find an infinite subsequence $\{a_{n_k}\}$ of the $\{a_n\}$ such that $a_{n_k} < a_{n_k+1}$ and $a_{n_k} \to \infty$, hence

$$\liminf \delta_n / \delta_{n-1} = 0.$$

Now if dy/dx exists, is finite, and is different from zero at the point x, δ_n/δ_{n-1} must tend to 1. Hence, at any point $x \in N$, dy/dx cannot exist, be finite, and be different from zero. But dy/dx exists and has a finite value almost everywhere. Then the only possible conclusion is $dy/dx = 0$ almost everywhere, which proves the singularity of the function.

Modulus of continuity of $?(x)$. We need the following result on continued fractions, which to our knowledge has not been stated:

LEMMA. *Let* $p_n/q_n = (0, a_1, a_2, \cdots, a_n)$. *Let* θ *be the Fibonacci number* $(1/2)(5^{1/2}+1)$. *We have the inequality* $q_n < \theta^{a_1+a_2+\cdots+a_n}$.

We shall prove this lemma by induction. We have $q_1 = a_1 < \theta^{a_1}$ for it is easily seen that $m < \theta^m$ for every positive integer m. We have also $q_0 = 1 = \theta^0$. And we have generally $q_k = a_k q_{k-1} + q_{k-2}$ $(k = 2, 3, \cdots, n)$. If supposing the lemma true for $n = k-2$ and $n = k-1$ we prove that it is true for $n = k$, we will have proved the result as stated. Let

$$q_{k-1} < \theta^{a_1 + \cdots + a_{k-1}}, \qquad q_{k-2} \leqq \theta^{a_1 + \cdots + a_{k-2}}.$$

It is sufficient to prove that

$$a_k \theta^{a_1 + \cdots + a_{k-1}} + \theta^{a_1 + \cdots + a_{k-2}} \leqq \theta^{a_1 + \cdots + a_k},$$

that is, $a_k \theta^{a_{k-1}} + 1 \leqq \theta^{a_{k-1} + a_k}$ or $a_k + 1/\theta^{a_{k-1}} \leqq \theta^{a_k}$. Hence it is sufficient to prove that $a_k + 1/\theta \leqq \theta^{a_k}$. Now for $a_k = 1$ we have the equality $1 + 1/\theta = \theta$ and it is easy to see that $2 + 1/\theta = \theta^2$ and that the function $\theta^x - x$ increases when $x \geqq 2$. Hence the lemma is proved. (It is easy to see, by considering the number $(0, 1, 1, 1, \cdots)$ that this result is the best possible of its kind, in order of magnitude.)

We can now proceed to determine the modulus of continuity of $?(x)$. In the definition of the function by successive approximations, we start from the

Farey fractions 0/1 and 1/1 and in a first operation we introduce the mediant 1/2, in a second operation the two mediants 1/3 and 2/3, in a third operation, four mediants, and so on. In the pth operation we introduce 2^{p-1} mediants and we get a sequence of fractions containing

$$2 + (1 + 2 + \cdots + 2^{p-1}) = 2^p + 1$$

fractions, which we can call the Minkowski sequence of order p and denote by $\mathfrak{M}_p$. To the sequence $\mathfrak{M}_p$ corresponds, by the transformation $y = ?(x)$, the sequence of numbers $k/2^p$ $(k = 0, 1, 2, \cdots, 2^p)$. The formula (11) giving the value of $?(x)$ shows that the fractions belonging to $\mathfrak{M}_p$ are those which, when written in the form $(0, a_1, a_2, \cdots, a_n)$, are such that $\sum_1^n a_\nu$ does not exceed $p+1$. Hence, by the lemma, if α/β belongs to $\mathfrak{M}_p$, we have $\beta < \theta^{p+1}$, and this order of magnitude is actually attained for the fraction $(0, 1, 1, \cdots, 1)$ where 1 is repeated $p+1$ times. Now it is immediately seen, by induction, that if α/β, α'/β' are two consecutive fractions of $\mathfrak{M}_p$, we have $|\beta'\alpha - \beta\alpha'| = 1$ and thus the distance between two consecutive fractions of $\mathfrak{M}_p$ is greater than $1/\theta^{2p+2}$.

Let now x, x' be two irrational points of $(0, 1)$, $y = ?(x)$, $y' = ?(x')$. At a certain stage of the dissection *one* fraction x_0 appears for the first time in (x, x'). Let us continue the dissection until one fraction appears for the first time in (x, x_0) or in (x_0, x') or in both intervals. Let this stage of the dissection be the pth, then we have $x' - x > 1/\theta^{2p+2}$ and $y' - y < 4/2^p$. Hence,

$$(2p + 2) \log \theta > \log \left| \frac{1}{x' - x} \right|, \qquad (p - 2) \log 2 < \log \left| \frac{1}{y' - y} \right|$$

which proves that

$$| y' - y | < | x' - x |^{2^{-1} \log 2 / \log \theta},$$

C being an absolute constant, and this relation being true for every couple of irrationals x, x' is also valid for x or x' or both rational. Hence: *the function* $?(x)$ *satisfies a Lipschitz condition of order* $\alpha = (1/2) \log 2 / \log \theta$ *where* θ *is the Fibonacci number* $(1/2)(5^{1/2} + 1)$.

We shall now prove that α is the best possible exponent for the Lipschitz condition of $?(x)$ and that it cannot be improved.

Let us consider, in fact, the number

$$\chi = (0, 1, 1, \cdots) = (5^{1/2} - 1)/2 = 1/\theta.$$

The corresponding value of the function is

$$\eta = ?(\chi) = 1 - 1/2 + 1/2^2 - \cdots = 2/3.$$

Let p_n/q_n be the successive convergents of χ. It is well known that

$$q_n = (1/5^{1/2}) [\theta^{n+1} - (-1)^{n+1}(1/\theta^{n+1})], \qquad p_n = q_{n-1}.$$

Now

$$\delta = |\chi - p_n/q_n| < 1/q_n^2$$

which is of the same order as $1/\theta^{2n}$, whereas

$$\eta - ?(p_n/q_n) = (-1)^n(1/2^n) + (-1)^{n+1}(1/2^{n+1}) + \cdots$$

is of order $1/2^n$, that is, of order $\delta^{2^{-1}\log 2/\log\theta}$, which proves that the number α of our Lipschitz condition is the best possible one.

Fourier-Stieltjes coefficients of $?(x/2\pi)$. Let

$$c_n = \int_0^{2\pi} e^{nix} d?(x/2\pi).$$

It is immediately seen that

$$c_n = \lim_{p=\infty}\left[\frac{1}{2^p}\sum_{\rho\subset\mathfrak{M}_p} e^{2\pi ni\rho}\right]$$

where the summation is extended to all fractions ρ belonging to $\mathfrak{M}_p$.

It does not seem to be known whether c_n tends to zero for $n=\infty$. If we confine ourselves to the behavior of c_n "in the average," we get the following result. It is well known by a theorem of Wiener(5) that

$$|c_1|^2 + |c_2|^2 + \cdots + |c_n|^2 < An\omega(1/n)$$

$\omega(\delta)$ being the modulus of continuity of the function and A an absolute constant. Hence, by our result on the modulus of continuity of $?(x)$ we have

$$|c_1|^2 + |c_2|^2 + \cdots + |c_n|^2 = O(n^{1-2^{-1}\log 2/\log\theta})$$

and, by Schwarz's inequality

$$|c_1| + |c_2| + \cdots + |c_n| = O(n^{1-4^{-1}\log 2/\log\theta}).$$

---

(5) See, for example, Zygmund, *Trigonometrical series*, p. 221.

Harvard University,
Cambridge, Mass.
Massachusetts Institute of Technology,
Cambridge, Mass.

ON THE PARTIAL SUMS OF FOURIER SERIES AT POINTS OF DISCONTINUITY

BY

OTTO SZÁSZ

1. **Introduction.** Consider a Fourier sine series

$$f(\theta) \sim \sum_1^\infty b_\nu \sin \nu\theta, \qquad b_\nu = (2/\pi)\int_0^\pi f(\theta) \sin \nu\theta d\theta, \qquad 0 < \theta < \pi, \tag{1.1}$$

and write

$$s_n(\theta) = \sum_1^n b_\nu \sin \nu\theta, \qquad n = 1, 2, 3, \cdots. \tag{1.2}$$

Fejér proved (cf. Zygmund [5, p. 181])([1]) that if $f(\theta)$ is of bounded variation, and if $n\theta_n \to \alpha$ as $\theta_n \to 0$, then

$$s_n(\theta_n) \to (2/\pi)f(+0)\int_0^\alpha \frac{\sin t}{t} dt \equiv (2/\pi)f(+0)I(\alpha). \tag{1.3}$$

In particular, choosing α so that $I(\alpha) = \pi/2 = \int_0^\infty t^{-1} \sin t dt$ (thus $\int_\alpha^\infty t^{-1} \sin t dt = 0$), we get $s_n(\theta_n) \to f(+0)$, which is half of the jump of $f(\theta)$ at $\theta = 0$.

On the other hand for $\alpha = \pi$, which gives $I(\alpha)$ its maximal value

$$s_n(\theta_n) \to (2/\pi)f(+0)\int_0^\pi \frac{\sin t}{t} dt = f(+0) \times 1.08949 \cdots.$$

Thus the limit points of the partial sums as $\theta_n \to 0$ cover an interval which extends beyond $f(+0)$, if $f(+0) \neq 0$. This is called Gibbs' phenomenon.

It was also proved by Fejér and Csillag (for references and further results see Szász [4]) that for functions of bounded variation

$$n^{-1}\sum_1^n \nu b_\nu \to (2/\pi)f(+0), \qquad \text{as } n \to \infty. \tag{1.4}$$

These facts suggest the consideration of

$$s_n(\theta_n) = \sum_1^n \nu b_\nu \frac{\sin \nu\theta_n}{\nu}, \qquad \theta_n \to 0,$$

as a transform of the sequence $\{nb_n\}$, that is, as a special case of the triangular type transform

Presented to the Society, September 8, 1942; received by the editors September 8, 1942.

([1]) Numbers in brackets refer to the literature at the end of this paper.

$$T_n = \sum_1^n a_{n\nu}\tau_\nu, \tag{1.5}$$

where now $\tau_\nu = \nu b_\nu$, $a_{n\nu} = \nu^{-1} \sin \nu\theta_n$. We shall not restrict ourselves to regularity conditions, and we shall not assume convergence of the sequence $\{\tau_n\}$, but merely Cesàro summability of some order. We then seek simple necessary and sufficient conditions for the convergence of the transform T_n (in general to a different limit). The application to Fourier sine series yields a generalized Gibbs' phenomenon, and also a new device to determine the generalized jump of a function. Our results are in close relationship with some results of Rogosinski [1, 2].

We consider more generally the transform

$$T(\rho_n, \theta_n) = \sum_1^n \tau_\nu \rho_n^\nu \nu^{-1} \sin \nu\theta_n, \qquad \rho_n \to 1, \theta_n \to 0, \tag{1.6}$$

which in the case $\tau_\nu = \nu b_\nu$ becomes $\sum_1^n \rho_n^\nu b_\nu \sin \nu\theta_n = s_n(\rho_n, \theta_n)$, where $s_n(\rho, \theta)$ is the nth partial sum of the harmonic series $\sum_1^\infty \rho^\nu b_\nu \sin \nu\theta$.

2. **Permanency with respect to convergent sequences.** It is well known that the convergence of the sequence $\{\tau_n\}$ implies the convergence of the transform T_n, if and only if

$$\lim_{n\to\infty} a_{n\nu} = 0, \qquad \text{for } \nu = 1, 2, 3, \cdots;$$

$$\sum_{\nu=1}^n | a_{n\nu} | = O(1), \qquad \text{as } n \to \infty;$$

$$\lim_{n\to\infty} \sum_1^n a_{n\nu} = \sigma \text{ exists.}$$

We then have $\lim T_n = \sigma \lim \tau_n$. If we restrict ourselves to sequences $\tau_n \to 0$, then the last condition can be omitted. Applied to (1.6) this yields the necessary and sufficient conditions:

$$\sum_1^n \rho_n^\nu \nu^{-1} | \sin \nu\theta_n | = O(1), \qquad \text{as } n \to \infty; \tag{2.1}$$

$$\lim \sum_1^n \rho_n^\nu \nu^{-1} \sin \nu\theta_n = \sigma. \tag{2.2}$$

In particular the last condition is $s_n(\rho_n, \theta_n) \to \sigma$ for the harmonic series $\sum_1^\infty \rho^\nu \nu^{-1} \sin \nu\theta = \arctan \{(\rho \sin \theta)/(1 - \rho \cos \theta)\}$.

We first assume

$$0 < \liminf \rho_n^n \leqq \limsup \rho_n^n < \infty; \tag{2.3}$$

in this case for some $c_1>0$, $c_2>0$

$$c_1\sum_1^n \nu^{-1}|\sin\nu\theta| < \sum_1^n \rho_n^\nu \nu^{-1}|\sin\nu\theta| < c_2\sum_1^n \nu^{-1}|\sin\nu\theta|,$$

thus (2.1) reduces to

$$\sum_1^n \nu^{-1}|\sin\nu\theta_n| = O(1), \qquad \text{as } n\to\infty. \tag{2.4}$$

Now for any $\theta>0$

$$\sum_1^n \nu^{-1}|\sin\nu\theta| < n\theta, \tag{2.5}$$

hence $n\theta_n=O(1)$ implies (2.4). To prove the converse let $\theta_n<1<\theta_n(n-1)$, and put $[\theta_n^{-1}]=\kappa_n=\kappa$, so that $\kappa\leqq\theta_n^{-1}<\kappa+1\leqq n$. Now $\sum_1^n\nu^{-1}|\sin\nu\theta_n| >(1/2)\sum_1^n\nu^{-1}(1-\cos 2\nu\theta_n)$, and

$$\left|\sum_1^n \nu^{-1}\cos 2\nu\theta_n\right| < \sum_1^\kappa \nu^{-1} + \left|\sum_{\kappa+1}^n \nu^{-1}\cos 2\nu\theta_n\right|$$

$$< 1+\log\kappa + (1/(\kappa+1))\max_{\kappa<\lambda\leqq n}\left|\sum_{\kappa+1}^\lambda \cos 2\nu\theta_n\right|$$

$$< 1+\log\theta_n^{-1} + \theta_n/\sin\theta_n < 3-\log\theta_n.$$

Thus

$$2\sum_1^n \nu^{-1}|\sin\nu\theta_n| > \log n+\log\theta_n - 3 = -3+\log(n\theta_n);$$

hence (2.4) implies $n\vartheta_n=O(1)$. For null sequences only this is required.

To satisfy (2.2) consider the case that 0 is a limit point of the sequence $\{n\theta_n\}$; for a subsequence of indices n: $n\theta_n\to 0$, and for that subsequence, using (2.5)

$$\sum_1^n \rho_n^\nu \nu^{-1}\sin\nu\theta_n = O\left(\sum_1^n \nu^{-1}|\sin\nu\theta_n|\right) = O(n\theta_n) = o(1).$$

Hence σ, if it exists, is 0 and then every convergent sequence is transformed into a null sequence. Next assume lim inf $n\theta_n>0$. We choose a subsequence of integers $n=n'$ for which ρ_n^n and $n\theta_n$ have limits $n'\theta_{n'}\to\beta>0$, $\rho_{n'}^{n'}\to e^\gamma$ say; by (2.3) γ is finite. Furthermore from $\log\rho/(\rho-1)\to 1$ as $\rho\to 1$, $n'(\rho_{n'}-1)\to\gamma$.

Suppose first $\gamma=0$, that is $\rho_{n'}^{n'}\to 1$, and $n'(\rho_{n'}-1)\to 0$. Now, as n runs through the sequence $\{n'\}$

$$\left|\sum_1^n (\rho_n^\nu-1)\nu^{-1}\sin\nu\theta_n\right| < |\rho_n-1|\,O(n)\sum_1^n \nu^{-1}|\sin\nu\theta_n| = o(1)O(n\theta_n) = o(1);$$

hence

$$\lim \sum_1^n \rho_n^\nu \nu^{-1} \sin \nu\theta_n = \lim \sum_1^n \nu^{-1} \sin \nu\theta_n$$

for $n = n' \to \infty$, if either side exists. But

$$\sum_1^n \nu^{-1} \sin \nu\theta = \int_0^\theta \left(\sum_1^n \cos \nu t \right) dt = -(1/2)\theta + \int_0^\theta \frac{\sin (n + 1/2)t}{2 \sin (t/2)} dt$$
$$= -(1/2)\theta + \int_0^{(n+1/2)\theta} \frac{\sin u du}{(2n+1) \sin (u/(2n+1))},$$

hence

$$\sum_1^n \nu^{-1} \sin \nu\theta_n = -(1/2)\theta_n + \int_0^\beta \frac{\sin u}{u} \cdot \frac{u du}{(2n+1) \sin (u/(2n+1))} + o(1)$$
$$\to \int_0^\beta \frac{\sin u}{u} du$$

as $n \to \infty$ through the sequence $\{n'\}$. The consideration of the case $\gamma \neq 0$ remains; we write

$$\sum_1^n \rho^\nu \nu^{-1} \sin \nu\theta$$
$$= \int_0^\theta \left(\sum_1^n \rho^\nu \cos \nu t \right) dt$$
$$= \int_0^\theta \frac{\cos t - \rho^2 + \rho^{n+2} \cos nt - \rho^{n+1} \cos (n+1)t}{1 - 2\rho \cos t + \rho^2} dt$$
$$= \int_0^\theta \frac{1-\rho^2-(1-\cos t)+\rho^{n+1}[\cos nt - \cos (n+1)t] - (1-\rho)\rho^{n+1} \cos nt}{(1-\rho)^2 + 2\rho(1-\cos t)} dt,$$

thus

$$\begin{aligned} \sum_1^n \rho^\nu \nu^{-1} \sin \nu\theta = {} & (1-\rho^2) \int_0^\theta \frac{dt}{(1-\rho)^2 + 4\rho \sin^2 (t/2)} \\ & - 2 \int_0^\theta \frac{\sin^2 (t/2) dt}{(1-\rho)^2 + 4\rho \sin^2 (t/2)} \\ & - (1-\rho)\rho^{n+1} \int_0^\theta \frac{\cos nt dt}{(1-\rho)^2 + 4\rho \sin^2 (t/2)} \\ & + 2\rho^{n+1} \int_0^\theta \frac{\sin (t/2) \sin (n + 1/2)t dt}{(1-\rho)^2 + 4\rho \sin^2 (t/2)}. \end{aligned} \tag{2.6}$$

Now

$$\int_0^{\theta_n} \frac{\sin^2 (t/2)dt}{(1-\rho_n)^2 + 4\rho_n \sin^2 (t/2)} < \frac{\theta_n}{4\rho_n} \to 0 \qquad \text{as } n \to \infty.$$

Next

$$(\rho_n^2 - 1)\int_0^{\theta_n} \frac{dt}{(1-\rho_n)^2 + 4\rho_n \sin^2 (t/2)}$$

$$= n(\rho_n - 1)(\rho_n + 1)\int_0^{n\theta_n} \frac{du}{n^2[(\rho_n - 1)^2 + 4\rho_n \sin^2 (u/2n)]}$$

$$\to 2\gamma \int_0^{\beta} \frac{du}{\gamma^2 + u^2} = 2 \arctan (\beta/\gamma).$$

Similarly

$$(\rho_n - 1)\rho_n^{n+1}\int_0^{\theta_n} \frac{\cos ntdt}{(1-\rho_n)^2 + 4\rho_n \sin^2 (t/2)}$$

$$= n(\rho_n - 1)\rho_n^{n+1}\int_0^{n\theta_n} \frac{\cos udu}{n^2[(\rho_n - 1)^2 + 4\rho_n \sin^2 (u/2n)]} \to \gamma e^{\gamma}\int_0^{\beta} \frac{\cos udu}{\gamma^2 + u^2},$$

and

$$\rho_n^{n+1}\int_0^{\theta_n} \frac{\sin (t/2) \sin (n + 1/2)tdt}{(1-\rho_n)^2 + 4\rho_n \sin^2 (t/2)}$$

$$= \rho_n^{n+1}\int_0^{(n+1/2)\theta_n} \frac{(2n+1) \sin \{u/(2n+1)\} \cdot \sin udu}{n(2n+1)[(\rho_n - 1)^2 + 4\rho_n \sin^2 \{u/(2n+1)\}]}$$

$$\to (1/2)e^{\gamma}\int_0^{\beta} \frac{u \sin udu}{\gamma^2 + u^2}.$$

Summarizing

$$\sum_1^n \rho_n^{\nu-1} \sin \nu\theta_n \to \int_0^{\beta} \frac{\gamma e^{\gamma} \cos u + e^{\gamma} u \sin u - 2\gamma}{\gamma^2 + u^2} du$$

$$= \int_0^{\beta/\gamma} \frac{e^{\gamma}(t \sin \gamma t + \cos \gamma t) - 2}{1 + t^2} dt.$$

The case (2.3) is now completely discussed. We next assume

$$\limsup_{n\to\infty} \rho_n^n = \infty,$$

so that for a subsequence $n = n' : \rho_{n'}^{n'} \to \infty$. We first prove that (2.1) implies $n'\theta_{n'} \to 0$. Otherwise for a subsequence n'' of $n' : n''\theta_{n''} \to \beta > 0$. For these indices

$$\sum_1^n \rho_n^\nu \nu^{-1} |\sin \nu\theta_n| > \sum_{\nu \leqq \alpha/\theta_n} \rho_n^\nu \nu^{-1} |\sin \nu\theta_n|,$$

where α is so chosen that $0<\alpha<\beta$ and $\alpha \leqq \pi/2$. Now

$$\sum_{\nu \leqq \alpha/\theta_n} \rho_n^\nu \nu^{-1} |\sin \nu\theta_n| > (2/\pi)\theta_n \sum \rho_n^\nu = \frac{2\rho_n\theta_n}{\pi} \frac{\rho_n^{[\alpha\theta_n^{-1}]} - 1}{\rho_n - 1},$$

hence (2.1) implies

$$\theta_n \rho_n^{\alpha n/2\beta} = O(\rho_n - 1),$$

which by virtue of $\log \rho_n/(\rho_n - 1) \to 1$ yields $n\theta_n = o(1)$. Furthermore

$$(2/\pi)\theta_n \sum \rho_n^\nu < \sum_1^n \rho_n^\nu \nu^{-1} |\sin \nu\theta_n| < \theta_n \sum_1^n \rho_n^\nu,$$

thus, if for a subsequence of indices $\rho_n^n \to \infty$, then for these indices (2.1) is equivalent to

$$\theta_n \rho_n^n = O(\rho_n - 1).$$

If this condition is satisfied, then in view of $n\theta_n \to 0$

$$0 < \theta_n \sum_1^n \rho_n^\nu - \sum_1^n \rho_n^\nu \nu^{-1} \sin \nu\theta_n = \theta_n \sum_1^n \rho_n^\nu \left(1 - \frac{\sin \nu\theta_n}{\nu\theta_n}\right)$$

$$< \theta_n \left(1 - \frac{\sin n\theta_n}{n\theta_n}\right) \sum_1^n \rho_n^\nu \to 0,$$

hence (2.2) holds if and only if $\lim \theta_n \rho_n^n/(\rho_n - 1)$ exists, which is then the value of σ.

Finally assume that $\liminf \rho_n^n = 0$; thus for a subsequence of indices $\rho_n^n \to 0$ $(\gamma = -\infty)$. If $n\theta_n = O(1)$, then $\sum_1^n \rho_n^\nu \nu^{-1} |\sin \nu\theta_n| = O(\sum_1^n \nu^{-1} |\sin \nu\theta_n|) = O(1)$, which is (2.1). If on the other hand for a subsequence $n\theta_n \to \infty$, then

$$\sum_1^n \rho_n^\nu \nu^{-1} |\sin \nu\theta_n| > \sum_{\theta_n^{-1} < \nu}^{\leqq n} \rho_n^\nu \nu^{-1} \frac{1 - \cos 2\nu\theta_n}{2};$$

but $\rho_n^\nu \nu^{-1} \downarrow$ as $\nu \uparrow$, hence for $\theta_n < 1$

$$\left| \sum_{\theta_n^{-1} < \nu}^{\leqq n} \rho_n^\nu \nu^{-1} \cos 2\nu\theta_n \right| < \rho_n^{1+[\theta_n^{-1}]} \frac{1}{1 + [\theta_n^{-1}]} \max_{\lambda \leqq n} \left| \sum_{\theta_n^{-1} < \nu}^{\leqq \lambda} \cos 2\nu\theta_n \right|$$

$$< \frac{\theta_n}{\sin \theta_n} < \pi/2.$$

Furthermore

$$\sum_{\theta_n^{-1}<\nu}^{\leq n} \rho_n^\nu \nu^{-1} = \sum_{\theta_n^{-1}<\nu}^{\leq n} \nu^{-1} \exp(\nu \log \rho_n) > \sum \int_\nu^{\nu+1} u^{-1} \exp(-u \log \rho_n^{-1}) du$$
$$= \int_{1+[\theta_n^{-1}]}^{n+1} u^{-1} \exp(-u \log \rho_n^{-1}) du = \int_{(1+[\theta_n^{-1}])\log \rho_n^{-1}}^{(n+1)\log \rho_n^{-1}} t^{-1} e^{-t} dt;$$

thus in this case (2.1) implies $\theta_n = O(\log 1/\rho_n)$, or $\theta_n = O(1-\rho_n)$. If this condition is satisfied, then

$$\sum_1^n \rho_n^\nu \nu^{-1} |\sin \nu\theta_n| = O\left(\theta_n \sum_1^n \rho_n^\nu\right) = O(\theta_n/(1-\rho_n)) = O(1),$$

hence (2.1) holds. To satisfy (2.2) now, we note that

$$\sum_{n+1}^\infty \rho_n^\nu \nu^{-1} \sin \nu\theta_n = O\left(\theta_n \frac{\rho_n^n}{1-\rho_n}\right) = O(\rho_n^n) = o(1),$$

hence (2.2) holds if and only if

$$\sum_1^\infty \rho_n^\nu \nu^{-1} \sin \nu\theta_n = \operatorname{arc\,tan} \frac{\rho_n \sin \theta_n}{1 - \rho_n \cos \theta_n}$$

has a limit, and σ is then this limit. But

$$\frac{\rho_n \sin \theta_n}{1-\rho_n \cos \theta_n} = \frac{\rho_n \sin \theta_n}{1-\rho_n + \rho_n(1-\cos \theta_n)} \sim \frac{\theta_n}{1-\rho_n} \cdot \frac{1}{1+O(1-\rho_n)} \sim \frac{\theta_n}{1-\rho_n},$$

hence σ exists, if and only if $\lim \theta_n/(1-\rho_n) = \delta < +\infty$. We then have $\sigma = \lim \operatorname{arc\,tan} \{\theta_n/(1-\rho_n)\} = \operatorname{arc\,tan} \delta$. To summarize our results put

(a)
$$\sigma(\beta, 0) = \int_0^\beta \frac{\sin u}{u} du,$$
$$\sigma(\beta, \gamma) = \int_0^\beta \frac{\gamma e^\gamma \cos u + e^\gamma \sin u - 2\gamma}{\gamma^2 + u^2} du, \text{ for finite } \gamma \neq 0,$$

(b)
$$\sigma(0, \infty) = \lim \frac{\theta_n \rho_n^n}{\rho_n - 1},$$

(c)
$$\sigma(\delta, -\infty) = \lim \operatorname{arc\,tan} \frac{\theta_n}{1-\rho_n} = \operatorname{arc\,tan} \delta < \pi/2.$$

We then have

THEOREM 1. *Necessary and sufficient conditions that for every convergent sequence $nb_n \to \tau$ the transform $\sum_1^n \rho_n^\nu b_\nu \sin \nu\theta_n$ has a limit, are that one of the following three cases holds:*

(a′) $n(\rho_n-1)\to\gamma$ *finite*, $n\theta_n\to\beta<\infty$,
(b′) $n(\rho_n-1)\to+\infty$, $\lim\theta_n\rho_n^n(\rho_n-1)^{-1}$ *exists*,
(c′) $n(\rho_n-1)\to-\infty$, $\lim\theta_n(1-\rho_n)^{-1}=\delta$ *exists*, $0\leqq\delta<\infty$.

The limit of the transform is then $\tau\sigma$, where σ is defined above for the respective cases. Different subsequences may belong to different cases (β,γ) if only the corresponding σ attain the same value, and with the restriction $n\theta_n=O(1)$ in case (a′).

3. Permanency with respect to (C,κ) summability. Given the sequence $\{\tau_n\}$, write

$$\overset{0}{\tau}_n=\tau_n,\qquad \overset{\kappa}{\tau}_n=\sum_{\nu=1}^{n}\overset{\kappa-1}{\tau}_\nu,\qquad n,\kappa=1,2,3,\cdots;$$

also

$$A_n^{\kappa}=C_{n+\kappa,n}=\frac{(\kappa+1)\cdots(\kappa+n)}{n!}\sim\frac{n^{\kappa}}{\kappa!},\qquad\text{as } n\to\infty. \tag{3.1}$$

The sequence $\{\tau_n\}$ is summable (C,κ) to the value τ, if $\overset{\kappa}{\tau}_n/A_n^{\kappa}\to\tau$ as $n\to\infty$; $(C,0)$ is evidently convergence.

We write

$$\Delta^0\tau_n=\tau_n,\qquad \Delta^1\tau_n=\Delta\tau_n=\tau_n-\tau_{n+1},\qquad \Delta^{\kappa}\tau_n=\Delta(\Delta^{\kappa-1}\tau_n);$$

then by induction

$$\Delta^{\kappa}\tau_n=\sum_{\nu=0}^{\kappa}(-1)^{\nu}C_{\kappa,\nu}\tau_{n+\nu},\qquad \kappa=0,1,2,\cdots. \tag{3.2}$$

Abel's transformation yields for finite sums

$$\sum_1^n\alpha_\nu\tau_\nu=\sum_1^n\overset{1}{\tau}_\nu\Delta\alpha_\nu=\sum_1^n\overset{2}{\tau}_\nu\Delta^2\alpha_\nu=\cdots,$$

where $\alpha_{n+1}=0$, $\alpha_{n+2}=0,\cdots$. Applying this to (1.5) we get

$$T_n=\sum_{\nu=1}^{n}\overset{\kappa}{\tau}_\nu\Delta^{\kappa}a_{n\nu},$$

where $a_{n\nu}=0$ for $\nu>n$. Thus the transform converges for every (C,κ) summable sequence if in addition to the conditions of §2

$$\sum_{\nu=1}^{n}A_\nu^{\kappa}|\Delta^{\kappa}a_{n\nu}|=O(1)\qquad\text{as } n\to\infty.$$

In particular for the transform (1.6) we have the conditions (2.1), (2.2) and

$$\sum_{\nu=1}^{n-\kappa}A_\nu^{\kappa}|\Delta^{\kappa}\rho_n^{\nu}\nu^{-1}\sin\nu\theta_n|+\sum_{n-\kappa+1}^{n}A_\nu^{\kappa}|\delta_\nu|=O(1)\qquad\text{as } n\to\infty, \tag{3.3}$$

where, from (3.2)

$$\delta_\lambda = \sum_{\nu=0}^{n-\lambda} (-1)^\nu C_{\kappa,\nu} \rho_n^{\lambda+\nu} \frac{\sin(\lambda+\nu)\theta_n}{\lambda+\nu}, \qquad n-\kappa < \lambda \leqq n.$$

We first consider $(C, 1)$ summability ($\kappa=1$). Now (3.3) becomes

$$\sum_1^{n-1} (\nu+1) \left| \Delta \rho_n^\nu \nu^{-1} \sin \nu\theta_n \right| + (n+1)\rho_n^n n^{-1} \left| \sin n\theta_n \right| = O(1),$$

or

$$(3.4)\quad \sum_1^{n-1} \nu \left| \rho_n^\nu \nu^{-1} \sin \nu\theta_n - \rho_n^{\nu+1}(\nu+1)^{-1} \sin(\nu+1)\theta_n \right| + \rho_n^n \left| \sin n\theta_n \right| = O(1).$$

We consider in succession the different cases of Theorem 1.

(a′) For a sequence of indices $n\theta_n \to \beta < \infty$, $n(\rho_n - 1) \to \gamma$ finite, that is, $\rho_n^n \to e^\gamma > 0$. Thus $\rho_n^n \sin n\theta_n$ is $O(1)$, and

$$\sum_1^{n-1} \nu\rho_n^\nu \left| \nu^{-1} \sin \nu\theta_n - \rho_n(\nu+1)^{-1} \sin(\nu+1)\theta_n \right|$$

$$\leqq \sum_1^{n-1} \nu\rho_n^\nu \left| \nu^{-1} \sin \nu\theta_n - (\nu+1)^{-1} \sin(\nu+1)\theta_n \right|$$

$$+ \left| 1-\rho_n \right| \sum_1^{n-1} \rho_n^\nu \left| \sin(\nu+1)\theta_n \right|.$$

Now

$$\left| 1-\rho_n \right| \sum_1^{n-1} \rho_n^\nu \left| \sin(\nu+1)\theta_n \right| < \left| 1-\rho_n \right| \sum_1^n \rho_n^\nu < \rho_n \left| 1-\rho_n^n \right| = O(1),$$

and

$$\nu \left| \nu^{-1} \sin \nu\theta_n - (\nu+1)^{-1} \sin(\nu+1)\theta_n \right|$$

$$= \left| (\nu+1)^{-1} \sin(\nu+1)\theta_n - 2\sin(1/2)\theta_n \cos((2\nu+1)/2)\theta_n \right| < 2\theta_n;$$

hence

$$\sum_1^{n-1} \nu\rho_n^\nu \left| \nu^{-1} \sin \nu\theta_n - (\nu+1)^{-1} \sin(\nu+1)\theta_n \right| < 2\theta_n \sum_1^n \rho_n^\nu = O(n\theta_n) = O(1).$$

Hence in this case no additional condition results.

(b′) $n(\rho_n - 1) \to +\infty$, $\theta_n \rho_n^n (\rho_n - 1)^{-1} \to \sigma$. Hence $n\theta_n \to 0$, and now $\rho_n^n \sin n\theta_n = O(1)$ is equivalent to $n\theta_n \rho_n^n = O(1)$. Thus $\theta_n \rho_n^n (\rho_n - 1)^{-1} = n\theta_n \rho_n^n n^{-1} (\rho_n - 1)^{-1} \to 0$, that is, $\sigma = 0$. Now

$$\left| 1-\rho_n \right| \sum_1^{n-1} \rho_n^\nu \left| \sin(\nu+1)\theta_n \right| = O\left[(\rho_n - 1)\theta_n \sum_1^n \nu\rho_n^\nu \right]$$

$$= O\left[\theta_n \rho_n^n (\rho_n - 1)^{-1} \right] = o(1);$$

furthermore

$$\theta_n \sum_1^n \rho_n^\nu = O[\theta_n \rho_n^n (\rho_n - 1)^{-1}] = o(1);$$

hence (3.4) holds. Finally:

(c′) If $\lim \theta_n/(1-\rho_n) < \infty$ exists, and $n(\rho_n - 1) \to -\infty$, that is, $\rho_n^n \to 0$, then $\rho < \rho_n^n \sin n\theta_n \to 0$, $|1-\rho_n| \sum_1^{n-1} \rho_n^\nu |\sin(\nu+1)\theta_n| < (1-\rho_n)\sum_0^\infty \rho_n^\nu = 1$, and

$$\theta_n \sum_1^n \rho_n^\nu < \frac{\theta_n}{1-\rho_n} = O(1).$$

No additional condition appears in this case. Summarizing, we have

THEOREM 2. *Necessary and sufficient conditions that when* $\lim n^{-1}\sum_1^n \nu b_\nu = \tau$ *exists the transform* $\sum_1^n \rho_n^\nu b_\nu \sin \nu\theta_n$ *has a limit*: $\tau\sigma$, *are either of the alternatives*:
(a″) $n(\rho_n - 1) \to \gamma$, *finite*, $n\theta_n \to \beta < \infty$,
(b″) $n(\rho_n - 1) \to +\infty$, $n\theta_n \rho_n^n = O(1)$,
(c″) $n(\rho_n - 1) \to -\infty$, $\lim \theta_n (1-\rho_n)^{-1}$ *exists*.

The value of σ is in the cases (a″) and (c″) given by (a) and (c). In case (b″) $\sigma = 0$. Different subsequences may belong to different cases if only σ has the same value, with the restriction $n\theta_n = O(1)$ in case (a″).

We now consider (C, κ) summability for $\kappa > 1$. First of all, to satisfy (3.3) we must have

$$n^\kappa \sum_{\nu=0}^m (-1)^\nu C_{\kappa,\nu} \rho_n^{n-m+\nu} \frac{\sin (n-m+\nu)\theta_n}{n-m+\nu} = O(1), \quad m = 0, 1, \cdots, \kappa - 1. \tag{3.5}$$

Or

$$n^{\kappa-1} \rho_n^n \sin n\theta_n = O(1),$$

$$n^\kappa \rho_n^n \left\{ \frac{\sin (n-1)\theta_n}{n-1} - \kappa\rho_n \frac{\sin n\theta_n}{n} \right\} = O(1),$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots ,$$

$$n^\kappa \rho_n^n \left\{ \frac{\sin (n-\kappa+1)\theta_n}{n-\kappa+1} - \cdots + (-1)^{\kappa-1} \kappa \rho_n^{\kappa-1} \frac{\sin n\theta_n}{n} \right\} = O(1).$$

This is equivalent to

$$\begin{aligned} n^{\kappa-1} \rho_n^n \sin n\theta_n &= O(1), \\ n^{\kappa-1} \rho_n^n \sin (n-1)\theta_n &= O(1), \\ \cdots\cdots\cdots\cdots &\cdots , \\ n^{\kappa-1} \rho_n^n \sin (n-\kappa+1)\theta_n &= O(1). \end{aligned} \tag{3.6}$$

In case (a'') the first condition becomes $\sin n\theta_n = O(n^{1-\kappa})$, as $n\to\infty$; in particular $\sin n\theta_n \to 0$, thus in view of (a'') $n\theta_n \to \lambda\pi$, λ a positive integer or zero. On putting $n\theta_n = \lambda\pi + \epsilon_n$, we get $\sin \epsilon_n = O(n^{1-\kappa})$, or $n\theta_n - \lambda\pi = O(n^{1-\kappa})$. From the second condition now $\cos n\theta_n \sin\theta_n = O(n^{1-\kappa})$, as $n\to\infty$, or $\lambda\pi + \epsilon_n = O(n^{2-\kappa})$; hence for $\kappa = 2$, (3.6) reduces to

$$n\theta_n = \lambda\pi + O(n^{-1}). \tag{3.7}$$

For $\kappa > 2$ we must have

$$n\theta_n - \lambda\pi = \epsilon_n = O(n^{1-\kappa}) \quad \text{and} \quad \lambda\pi + \epsilon_n = O(n^{2-\kappa}),$$

hence $\lambda = 0$, and

$$n\theta_n = O(n^{1-\kappa}). \tag{3.7'}$$

It then follows that

$$n^{\kappa-1}\sin(n-\nu)\theta_n = O(1) \qquad \text{for } \nu = 0, 1, \cdots, \kappa - 1.$$

Furthermore, for the rest of (3.3)

$$\sum_1^{n-\kappa} A_\nu^\kappa \left| \Delta^\kappa \rho_n^\nu \nu^{-1} \sin \nu\theta_n \right| = O\left(\sum_1^{n-\kappa} \nu^\kappa \left| \Delta^\kappa \rho_n^\nu \nu^{-1} \sin \nu\theta_n \right| \right).$$

Now

$$\Delta^\kappa \rho^\nu \nu^{-1} \sin \nu\theta = \Delta^\kappa \rho^\nu \int_0^\theta \cos \nu t dt = R \int_0^\theta \Delta^\kappa z^\nu dt, \qquad z = \rho e^{it},$$

and, using (3.2)

$$\Delta^\kappa \rho^\nu \nu^{-1} \sin \nu\theta = R \int_0^\theta \sum_{\lambda=0}^{\kappa} (-1)^\lambda C_{\kappa,\lambda} z^{\nu+\lambda} dt = R \int_0^\theta z^\nu (1-z)^\kappa dt,$$

hence

$$\begin{aligned} \left| \Delta^\kappa \rho^\nu \nu^{-1} \sin \nu\theta \right| &< \rho^\nu \int_0^\theta \left| 1 - \rho e^{it} \right|^\kappa dt < \rho^\nu \int_0^\theta \left\{ (1-\rho)^2 + \rho t^2 \right\}^{\kappa/2} dt \\ &< \theta \rho^\nu \left\{ (1-\rho)^2 + \rho\theta^2 \right\}^{\kappa/2}. \end{aligned}$$

Thus

$$\begin{aligned} \sum_1^{n-\kappa} \nu^\kappa \left| \Delta^\kappa \rho_n^\nu \nu^{-1} \sin \nu\theta_n \right| &< \left(\sum_1^n \nu^\kappa \rho_n^\nu \right) \left\{ (1-\rho_n)^2 + \rho_n \theta_n^2 \right\}^{\kappa/2} \theta_n \\ &< \left\{ n^2 (1-\rho_n)^2 + \rho_n n^2 \theta_n^2 \right\}^{\kappa/2} \theta_n \sum_1^n \rho_n^\nu \\ &= O\left(\theta_n \sum_1^n \rho_n^\nu \right), \end{aligned} \tag{3.8}$$

and, from $\rho_n^n = O(1)$,

$$\theta_n \sum_1^n \rho_n^\nu = O(n\theta_n) = O(1).$$

Hence in case (a″) the additional condition is (3.7) for $\kappa=2$, and (3.7′) for $\kappa>2$.

In case (b″): $n(\rho_n-1)\to+\infty$, $n\theta_n\rho_n^n=O(1)$, as $n\to\infty$; hence $\rho_n^n\to+\infty$, and $n\theta_n\to 0$. Now (3.6) becomes

$$n^\kappa \rho_n^n \theta_n = O(1). \tag{3.9}$$

For large n evidently $\rho_n>1$, and

$$\theta_n \sum_1^n \rho_n^\nu < n\theta_n\rho_n^n = O(1)$$

(from (3.9)). In view of (3.8) now (3.3) holds. Thus in this case the additional condition is (3.9) (for $\kappa\geqq 2$).

Finally, in case (c″): $n(\rho_n-1)\to-\infty$ (that is $\rho_n^n\to 0$), and $\lim \theta_n/(1-\rho_n)=\delta<\infty$ exists. Now $\rho_n^n<1/(n(1-\rho_n))$, hence $n\theta_n\rho_n^n=O(1)$; thus for $\kappa=2$ condition (3.6) reduces to $n\rho_n^n \sin n\theta_n=O(1)$. While for $\kappa>2$ (3.6) reduces to $n^{\kappa-1}\rho_n^n \sin n\theta_n=O(1)$ and $n^{\kappa-1}\theta_n\rho_n^n=O(1)$. Furthermore, as $\rho_n<1$, $\theta_n\sum_0^n\rho_n^\nu<\theta_n/(1-\rho_n)=O(1)$, hence, in view of (3.8) now (3.3) is satisfied.

We summarize our results in

THEOREM 3. *In order that* $\lim \sum_1^n \rho_n^\nu b_\nu \sin \nu\theta_n=\tau\sigma$ *exists, whenever* (C,κ) $\lim nb_n=\tau$ *for some* $\kappa\geqq 2$, *necessary and sufficient conditions are the alternatives*:

(a‴) $n(\rho_n-1)\to\gamma$, *finite, and for* $\kappa=2$: $n\theta_n=\lambda\pi+O(n^{-1})$, λ *an integer, for* $\kappa>2$: $\theta_n=O(n^{-\kappa})$;

(b‴) $n(\rho_n-1)\to+\infty$, $n^\kappa\rho_n^n\theta_n=O(1)$;

(c‴) $n(\rho_n-1)\to-\infty$, $\lim \theta_n/(1-\rho_n)=\delta<\infty$ *exists, and for* $\kappa=2$: $n\rho_n^n \sin n\theta_n=O(1)$, *for* $\kappa>2$: $n^{\kappa-1}\rho_n^n(\theta_n+|\sin n\theta_n|)=O(1)$.

The value of σ is given in case (a‴) by (a), where for $\kappa=2$: $\beta=\lambda\pi$, for $\kappa>2$: $\beta=0$, $\sigma=0$; in case (b‴): $\sigma=0$; in case (c‴): $\sigma=\arctan\delta$.

4. **Application to Fourier series.** First consider a function of bounded variation and its Fourier sine series (1.1). It follows from the introduction that $\lim nb_n$, if it exists, is $(2/\pi)f(+0)$. Under the assumptions of Theorem 1 on ρ_n and θ_n, $s_n(\rho_n,\theta_n)\to\tau\sigma=(2\sigma/\pi)f(+0)$. In particular whenever $\sigma>\pi/2$, then we have an analogue of Gibbs' phenomenon. It is known that for functions of bounded variation

$$(1/n)\sum_1^n \nu b_\nu \to (2/\pi)f(+0);$$

more generally if (cf. Szász [3, Lemma 6])

$$(4.1) \qquad 2f_1(\theta) = (2/\theta)\int_0^\theta f(t)dt \to j, \qquad \text{as } \theta \downarrow 0,$$

and

$$(4.2) \qquad \lim_{\delta\downarrow 0}\ \liminf_{n\to\infty}\ \min_{0<\kappa<\delta_n} \sum_n^{n+\kappa} b_\nu \geqq 0,$$

then

$$(1/n)\sum_1^n \nu b_\nu \to j/\pi.$$

Hence, applying Theorem 2 we have

$$s_n(\rho_n, \theta_n) \to (j/\pi)\sigma(\beta, \gamma), \qquad \text{as } n(\rho_n - 1) \to \gamma \text{ and } n\theta_n \to \beta;$$

j is the generalized jump of $f(\theta)$ at $\theta = 0$. For $\gamma = 0$ this yields a generalization of formula (1.4). Note that

$$f_1(\theta) = \theta^{-1}\int_0^\theta f(t)dt = \sum_1^\infty b_\nu \frac{1-\cos\nu\theta}{\nu\theta} = (\theta/2)\sum_1^\infty \nu b_\nu\left(\frac{\sin(\nu\theta/2)}{\nu\theta/2}\right)^2;$$

$(2\theta/\pi)\{s_0/2 + \sum_1^\infty((\sin\nu\theta)/\nu\theta)^2 s_\nu\}$ is called the Riemannian mean of the second kind corresponding to the sequence $\{s_n\}$. It is a regular transform, as is seen from the identity

$$\frac{2\theta}{\pi}\left\{1/2 + \sum_1^\infty\left(\frac{\sin\nu\theta}{\nu\theta}\right)^2\right\} = 1.$$

If we assume only that $(C, 2)\lim nb_n = j/\pi$ exists, then Theorem 3 yields again a Gibbs' phenomenon in the case (a''') and $\lambda > 0$.

In this connection we introduce two lemmas.

LEMMA 1. *If*

$$(4.3) \qquad (1-r)\sum_1^\infty \tau_n r^n \to \tau \qquad \text{as } r \uparrow 1,$$

and

$$(4.4) \qquad \tau_n' = \sum_1^n \tau_\nu > -pn,$$

for some $p > 0$, and all $n > 0$, then

$$(4.5) \qquad (C, 2)\lim \tau_n = \tau.$$

We have from (4.3)

$$(1-r)^2\sum_1^\infty \tau_n' r^n \to \tau \qquad \text{as } r \uparrow 1,$$

hence

$$(1-r)^2\sum_1^\infty (\tau_n' + pn)r^n \to \tau + p \qquad \text{as } r \uparrow 1;$$

in view of (4.4) a theorem of Hardy and Littlewood yields

$$\sum_1^n (\tau_\nu' + p\nu) \sim (1/2)(\tau + p)n^2,$$

or

$$\sum_1^n \tau_\nu' \sim (1/2)\tau n^2, \qquad \text{as } n \to \infty,$$

which is (4.5).

LEMMA 2. *If* (4.1) *holds, then* $(1-r)\sum_1^\infty nb_n r^n \to j/\pi$. [3, Lemma 5].

Combining these two lemmas it is seen that (4.1) and the assumption

$$\sum_1^n \nu b_\nu > -pn \qquad \text{for some } p > 0 \text{ and all } n > 0, \tag{4.6}$$

imply $(C, 2)$ $\lim nb_n = j/\pi$. With reference to Theorem 3 the assumptions (4.1) and (4.6) again yield a Gibbs' phenomenon.

In closing we remark that the existence of $f(+0)$ implies itself $(C, 2)$ $\lim nb_n = (1/2)f(+0)$. A more general result will be given elsewhere.

REFERENCES

1. W. Rogosinski, *Ueber den Einfluss einseitiger Eigenschaften einer Funktion auf ihre Fourierreihe*, Schriften der Koenigsberger Gelehrten Gesellschaft Naturwissenschaftliche Klasse vol. 3 (1926) pp. 57–98.

2. ———, *Abschnittsverhalten bei trigonometrischen und insbesondere Fourierschen Reihen*, Math. Zeit. vol. 41 (1936) pp. 75–136.

3. O. Szász, *Convergence properties of Fourier series*, Trans. Amer. Math. Soc. vol. 37 (1935) pp. 483–500.

4. ———, *The jump of a function determined by its Fourier coefficients*, Duke Math. J. vol. 4 (1938) pp. 401–407.

5. A. Zygmund, *Trigonometrical series*, 1935.

UNIVERSITY OF CINCINNATI,
CINCINNATI, OHIO

THE CHARACTERISTIC OF A QUADRATIC FORM FOR AN ARBITRARY FIELD

BY

RUFUS OLDENBURGER

1. **Introduction.** Ernst Witt [1](1) has shown that for a field K with characteristic not 2 each quadratic form Q can be transformed into a form $G+H$ where

$$G = \sum_{i=1}^{\sigma} (x_i^2 - y_i^2),$$

H is a nonzero form, that is, does not represent zero properly, the rank of Q is the sum of the ranks of G and H, and G has rank 2σ. The number σ is an invariant of Q under nonsingular linear transformations on Q. For the real field the number σ was defined by Loewy [2] as the minimum of the indices(2) of Q and $-Q$, and termed the *characteristic* of Q. Loewy showed that this characteristic could be defined in terms of exponents of elementary divisors of pencils $\{\rho F - Q\}$ formed from Q and real quadratic forms $\{F\}$. The definition of Loewy does not extend to an arbitrary field K, whereas the characteristic of Q for K is arrived at by Witt through the examination of a sequence of quadratic forms $Q, H_1, \cdots, H_\sigma$, where for each s the form Q is $G_s + H_s$ for

$$G_s = \sum_{i=1}^{s} (L_i^2 - M_i^2),$$

the L's and M's being linearly independent linear forms, while the rank of Q is the sum of the ranks of G_s and H_s. In the present paper we shall show that the characteristic of Q can be defined in terms of linearly independent linear forms directly associated with Q. This definition is particularly convenient for treating sums of forms.

With the aid of the viewpoint developed here it is proved (§§3–4) that the characteristic of a quadratic form Q is related to the ranks of the principal minors of matrices associated with Q. By means of this relation we are able to treat the characteristics of sums of forms, and to show in particular (§5) that the characteristic of $Q+\lambda L^2$, L linear, differs at most by 1 from that of Q. This property is likewise possessed by the rank of Q, and, if K is real, also by the index of Q. The above result on the characteristic of $Q+\lambda L^2$ will be used in another paper [3] to show that the characteristic σ of a quadratic

Presented to the Society, April 3, 1942; received by the editors February 13 and September 2, 1942.

(1) The numbers in brackets refer to the bibliography at the end of the paper.

(2) The index of Q is the number h of $+$ signs in a canonical form $x_1^2 + \cdots + x_h^2 - x_{h+1}^2 - \cdots - x_r^2$ to which Q is equivalent under a nonsingular linear transformation.

form Q determines the minimum value τ for which Q can be written as

$$\sum_{i=1}^{\tau} L_i M_i$$

where the L's and M's are linear forms. The maximum value which can be attained by the characteristic of Q relative to the rank r of Q is $[r/2]$, where $[r/2]$ designates the largest integer not exceeding $r/2$.

A sum

$$Q = \sum_{i=1}^{r} \lambda_i L_i^2,$$

where the L's are linear, and r is the rank of Q, is a *minimal representation* of Q.

2. **Preliminary definitions and conventions.** Throughout the present paper the usual restriction that the *characteristic* of the field K be different from 2 is made in order that each quadratic form Q may be written as

$$Q = \sum_{i,j=1}^{n} a_{ij} x_i x_j \tag{2.1}$$

where the matrix (a_{ij}) of coefficients is symmetric. We term (a_{ij}) the *matrix* A of Q.

In what follows we shall use the term "equivalent" to mean equivalent under nonsingular linear transformations.

We define the *characteristic* of a quadratic form Q to be the maximum number σ of linearly independent linear forms $L_1, \cdots, L_\sigma$ such that the rank of

$$Q + \lambda_1 L_1^2 + \cdots + \lambda_\sigma L_\sigma^2 \tag{2.2}$$

is identical with the rank of Q for all values of the λ's. In what follows the term "characteristic" will be understood to refer to the invariant just defined, until we have proved (§3) that this invariant is identical with the characteristic of Witt, and for the real field with that of Loewy.

If a form Q with rank r is written as a sum $G+H$ where G has rank 2σ and characteristic σ, and H has rank $r-2\sigma$ and characteristic 0, we term $G+H$ a *characteristic splitting* of Q. If we write Q as

$$\sum_{i=1}^{\sigma} L_i M_i + H, \tag{2.3}$$

where the component in the L's and M's is identical with the form G in a characteristic splitting of Q, we have a *decomposition* of Q.

If the rank of the form (2.2) is identical with the rank r of Q for all values of the λ's, the forms $L_1, \cdots, L_\sigma$ clearly depend only on the variables which occur in Q. A stronger statement can be made. We let

$$\mu_1 M_1^2 + \cdots + \mu_r M_r^2$$

be a minimal representation of Q. Since $M_1, \cdots, M_r$ may be taken as the variables in terms of which Q is expressed, it follows that the L's are linear forms in the M's, whence the characteristic of Q cannot exceed r.

3. **Characteristic splittings.** In the present section we shall relate the characteristic of a quadratic form Q to the results of Witt.

We recall that a matrix D of order c and rank d has ***nullity*** $c-d$.

LEMMA 3.1. *The characteristic of a quadratic form Q of rank r is the maximum σ for which Q is equivalent to a form F with rth order matrix C where complementary principal minors C_{11}, C_{22} of C have order and nullity σ, respectively.*

In considering matrices $\{C\}$ with complementary principal minors C_{11} and C_{22}, it will be no restriction if we take C_{11} to be a leading minor so that

$$C = \begin{Vmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{Vmatrix}. \tag{3.1}$$

We write Q as in (2.1) with $n=r$. We suppose that Q has characteristic σ. We may assume without restriction that the rank of (2.2) with $L_i \equiv x_i$ for each i is identically equal to r for all values of the λ's. We assume that $\sigma \geqq 1$. We let M denote the minor of the matrix A of Q obtained from A by deleting the first σ rows and σ columns of A. Expanding the determinant of the form (2.2) we find that the rank of the form (2.2) is r for all values of the λ's if and only if each principal minor determinant of A containing M vanishes, except the determinant $|A|$ itself. The rank of M is a number b, where $b < r-\sigma$. We may suppose that M is in the shape

$$\begin{Vmatrix} D & 0 \\ 0 & 0 \end{Vmatrix} \tag{3.2}$$

where D is a nonsingular minor of order b. We suppose, for the moment, that $b \geqq r-2\sigma$, and consider the minor determinants of A of the type

$$\begin{vmatrix} * & * & E' \\ * & D & 0 \\ E & 0 & 0 \end{vmatrix}, \tag{3.3}$$

where E has order $r-\sigma-b$, and E' denotes the transpose of E. If $b > r-2\sigma$, the minors of type (3.3) are distinct from $|A|$. Since the vanishing (3.3) implies the singularity of E, the last $r-\sigma-b$ columns of A are linearly dependent, a contradiction. Thus $b \leqq r-2\sigma$. If $b < r-2\sigma$, the matrix A is singular. Thus $b = r-2\sigma$.

Conversely, if Q has the matrix C where C_{22} has rank $r-2\sigma$ and order $r-\sigma$, we adjoin elements from q rows and q columns of C to C_{22} to obtain a minor

of C of order m, where $m=r-\sigma+q$, with rank at most r', where $r'=r-2\sigma+2q$. If $\sigma>0$, the relation $q<\sigma$ implies that $r'<m$, whence each square minor of C containing C_{22}, except C, is singular. The rank of (2.2) with $L_i\equiv x_i$ for each i is now r for all choices of the λ's, whence the characteristic of Q is at least σ.

THEOREM 3.1. *A quadratic form Q with rank r has characteristic σ if and only if Q has the characteristic splitting $G+H$, where G has rank 2σ and characteristic σ, while H has rank $r-2\sigma$ and characteristic* 0.

By a result of Witt, quoted in the introduction, the form Q is equivalent to a sum

$$\sum_{i=1}^{\rho} u_i v_i + H, \tag{3.4}$$

where H is a nonzero form with rank $r-2\rho$. The number ρ (by the theory of Witt) is uniquely determined by Q. Further, if Q is equivalent to a form (3.4) where H is a zero form with rank $r-2\rho$, the form Q is equivalent to a sum (3.4) with ρ replaced by a larger number ρ', where H now is a nonzero form with rank $r-2\rho'$.

We let σ denote the characteristic of Q. Since the rank of

$$Q+\lambda_1 u_1^2+\cdots+\lambda_\rho u_\rho^2$$

is r for all values of the λ's we have $\sigma\geqq\rho$.

By Lemma 3.1 the form Q is equivalent to a form Q' with the matrix C, given in (3.1), where the order of C_{11} and the nullity of C_{22} equal σ. In view of the nullity of C_{22} we may assume that C_{22} is written as (3.2) where the order of D is equal to $r-2\sigma$. We write Q' as (2.1) with $n=r$, whence $(a_{ij})=C$. The form Q' is now the sum $G'+H'$, where

$$G'=\sum_{i=1}^{\sigma} x_i L_i, \tag{3.5}$$

the L's are linear forms and H' is the form with matrix D. Since Q' has rank r, the variables $x_1,\cdots,x_\sigma$ in G', as well as $x_{\sigma+1},\cdots,x_{r-\sigma}$ in H', and $L_1,\cdots,L_\sigma$ comprise a set of linearly independent forms, whence these may be taken as the variables in terms of which Q' is expressed. It follows from the Witt theory that $\rho\geqq\sigma$, whence $\rho=\sigma$.

It is readily seen that the component in the x's and L's in (3.5) has characteristic σ, whereas H' has characteristic 0.

From Theorem 3.1 there are a number of immediate consequences valid for an arbitrary field K. The characteristic of a quadratic form Q of rank r does not exceed $[r/2]$. The *characteristic* of a quadratic form Q of rank r attains the *maximum value* $r/2$ if and only if Q is equivalent to the canonical form

$$\sum_{i=1}^{r/2} x_i y_i.$$

A quadratic form Q is a *nonzero form if and only if* the *characteristic* of Q is 0.

For the complex field the characteristic of a quadratic form Q is $[r/2]$. For the real field the characteristic is clearly the minimum of the indices of Q and $-Q$. The index and characteristic of Q may thus be distinct. In fact the concept of index of Q is identical with that of characteristic and the type $(\pm)$ of definiteness of a nonzero component H in a canonical splitting of Q.

From a result of Dickson [4] one can readily show that the quadratic forms

$$Q = \sum_{i=1}^{n} \alpha_i x_i^2, \qquad E = \sum_{i=1}^{n} e_i y_i^2,$$

with $\alpha_1 = e_1$ and rank n, are equivalent if and only if the subforms

$$\sum_{i=2}^{n} \alpha_i x_i^2, \qquad \sum_{i=2}^{n} e_i y_i^2$$

are equivalent. Witt [1] proved this same result by different methods, and showed that quadratic forms with canonical splittings $G+H$ and $G'+H'$, where H and H' are nonzero forms, are equivalent if and only if they have the same characteristic, and H is equivalent to H'. Thus H is uniquely determined up to equivalence. The study of the equivalence of quadratic forms thus reduces to that of nonzero forms, so extensively treated in the literature([3]).

4. **Characteristics and principal minors.** To treat characteristics of sums of forms we shall need some properties of principal minors developed here.

LEMMA 4.1. *If the matrix of order r of a quadratic form Q of rank r has a principal minor of nullity σ, the characteristic of Q is at least σ.*

Without restriction on the generality of the method we may suppose that the matrix of order r of $Q(x_1, \cdots, x_r)$ is given by

$$(4.1) \qquad A = \begin{Vmatrix} D_{11} & D_{12} & D_{13} \\ D_{21} & D_{22} & 0 \\ D_{31} & 0 & 0 \end{Vmatrix},$$

where the bottom right zero represents a minor of order σ, and D_{22} is nonsingular. We let t designate the number of rows of D_{13}. By Lemma 3.1 we may restrict ourselves to the case where $t > \sigma$. By a nonsingular linear transformation affecting only the variables $x_1, \cdots, x_t$, the form Q can be brought into a form Q' with matrix B, where the principal minor of B obtained by

([3]) See, for example, [5].

striking out the first t rows and columns of B is identical with this minor for A, and D_{12} is replaced by a nonsingular σth order minor followed by rows of zeros. The matrix B is of the type C given in (3.1), where C_{11}, C_{22} have order and nullity σ, respectively, whence by Lemma 3.1 the characteristic of Q is at least σ.

In Lemma 3.1 we showed how the characteristic σ of a quadratic form Q is the maximum value σ for which certain minors C_{11}, C_{22} possess given properties. We shall show how a further examination of these minors reveals whether or not the maximum value σ is attained for them. In the following theorem the characteristic of C_{22} is understood to be the characteristic of the quadratic form associated with C_{22}.

THEOREM 4.1. *Suppose that the order and nullity of complementary principal minors C_{11}, C_{22} of the matrix of order r of a quadratic form Q of rank r equal σ, respectively. The characteristic of Q is σ if and only if the characteristic of C_{22} is* 0.

We write Q as in (2.1) with $n=r$. Since C_{22} has nullity σ there is a nonsingular matrix M such that $MC_{22}M'$ is the minor (3.2) where D has order $r-2\sigma$. It will thus be no restriction on the generality of the method to suppose that the matrix A of Q has the shape (4.1) with $D_{22}=D$, and the order of D_{11} equal to σ. We can thus write Q as a sum $G'+H'$, where G' is given by (3.5), H' is a form with the matrix D, and the rank of Q is the sum of the ranks of G' and H'. The characteristic of G' is clearly σ.

If Q has characteristic σ, the sum $G'+H'$ is a characteristic splitting of Q, whence H' has characteristic 0. It follows that C_{22} has characteristic 0.

If conversely, the characteristic of C_{22} is 0, the sum $G'+H'$ is again a canonical splitting, whence the characteristic of Q is σ.

We consider the canonical splitting $G+H$, where

$$G = \sum_{i=1}^{\sigma} x_i x_{r-\sigma+i},$$

and H is a form in $x_{\sigma+1}, \cdots, x_{r-\sigma}$. Since the matrix of $G+H$ is of the type (3.1) with the order of C_{11} and the nullity of C_{22} equal to σ, and the characteristic of C_{22} equal to 0, the characteristic of a quadratic form Q is σ *if and only if* Q is equivalent to a form with the matrix (3.1) where complementary principal minors C_{11} and C_{22} have the properties just mentioned.

5. **The characteristic of a sum of forms.** Each quadratic form Q is equivalent to a quadratic form with a diagonal matrix. This is the same as the property that each quadratic form Q has a minimal representation. It follows that the study of the effect of the addition of a quadratic form F to a quadratic form Q reduces to the study of the addition of a term λL^2, L linear, to Q.

THEOREM 5.1. *Under addition of a term λL^2, L linear, to a quadratic form Q the characteristic σ of Q changes at most by* 1.

We suppose that the characteristic of $Q+\lambda L^2$ is at least $\sigma+2$. We let q designate the rank of $Q+\lambda L^2$. By Lemma 3.1 the pair (Q, L) can be transformed nonsingularly into a pair (Q', M), where the matrix C of $Q'+\lambda M^2$ is of order q, and is of the shape (3.1), while C_{11} and C_{22} have order and nullity $\sigma+2$, respectively.

The rank of a quadratic form Q changes at most by 1 under addition of a term λL^2, L linear, to Q. We shall assume, to begin with, that $q=r$, where r is the rank of Q. The addition of $-\lambda M^2$ to $Q'+\lambda M^2$ changes the nullity of C_{22} by at most 1. Thus C_{22} goes into a minor D_{22} with nullity at least $\sigma+1$. By Lemma 4.1 the characteristic of Q' is at least $\sigma+1$, a contradiction. It follows that the characteristic t of $Q+\lambda L^2$ does not exceed $\sigma+1$. Thus

$$\sigma + 1 \geqq t \geqq \sigma - 1.$$

We now consider the case where $q=r+1$, and designate the variables in $Q'+\lambda M^2$ by $y_1, \cdots, y_{r+1}$ where the rank of

$$Q' + \lambda M^2 + \lambda_1 y_1^2 + \cdots + \lambda_{\sigma+2} y_{\sigma+2}^2$$

is $r+1$ for all values of $\lambda_1, \cdots, \lambda_{\sigma+2}$. By the development used in the proof of Lemma 3.1 the form $Q'+\lambda M^2$ has a matrix C of order $r+1$ as given in (3.1) where C_{11} and C_{22} are minors of order and nullity $\sigma+2$, respectively. If M is linearly independent of $y_1, \cdots, y_{\sigma+2}$, we may suppose that $M\equiv y_{\sigma+3}$. Removal of the row and column of C corresponding to $y_{\sigma+3}$ yields the matrix of Q' of order r. The minor obtained from C_{22} by removal of this row and column has nullity at least $\sigma+1$. By Lemma 4.1 the characteristic of Q' is at least $\sigma+1$, a contradiction. If, on the other hand, the form M is linearly dependent on $y_1, \cdots, y_{\sigma+2}$, the minor C_{22} of C is a minor of a matrix C^* of order $r+1$ of Q'. Since C_{22} has nullity $\sigma+2$, there is a nonsingular matrix N such that $N'C^*N=A$, where A is given in (4.1) the minors D_{13} and D_{31} being of order $\sigma+2$, while D_{22} is nonsingular. Since A is singular, D_{13} is singular. It follows that there is a nonsingular matrix M such that $M'AM$ is identical in shape with A except that the last column of D_{13} is replaced by a column of zeros, and a corresponding remark holds for the last row of D_{31}. We drop the last row and column of $M'AM$ to obtain a matrix B whose lower principal minor of order $r-\sigma-3$ has nullity $\sigma+1$. By Lemma 4.1 and the invariance of the characteristic under nonsingular linear transformations we have again arrived at a contradiction. Thus in any event, when $q=r+1$, the characteristic of $Q+\lambda L^2$ does not exceed $\sigma+1$.

Since the index of Q is σ there are linearly independent linear forms $L_1, \cdots, L_\sigma$ such that the rank of (2.2) is r for all values of the λ's. If $q=r+1$, the form L is linearly independent of the variables in Q, whence the rank of

$$Q + \lambda L^2 + \lambda_1 L_1^2 + \cdots + \lambda_\sigma L_\sigma^2$$

is $r+1$ for all values of $\lambda_1, \cdots, \lambda_r$. Thus if the rank of $Q+\lambda L^2$ exceeds the rank of Q the characteristic of $Q+\lambda L^2$ is at least as great as that of Q.

The case where $q=r-1$ reverts to the preceding.

COROLLARY 5.1. *If ρ and R denote the characteristic and rank of the quadratic forms Q and F, respectively, the characteristic of $Q+F$ satisfies the inequalities:*

$$\rho - R \leqq \sigma \leqq \rho + R.$$

We shall show that Theorem 5.1 is valid if we replace the characteristic of Q by the index of Q for the real field. In the following theorem all coefficients are understood to be in the real field.

THEOREM 5.2 (Analogue of Theorem 5.1). *Under addition of a term λL^2, L linear, to a real quadratic form Q, the index of Q changes at most by* 1.

It is readily seen that Theorem 5.2 is true when the ranks of Q and $Q+\lambda L^2$ are distinct. We suppose, therefore, that these ranks are identical. We have

$$Q + \lambda L^2 \equiv M_1^2 + \cdots + M_q^2 - M_{q+1}^2 - \cdots - M_r^2,$$

where r is the rank of Q, while q is the index of $Q+\lambda L^2$, and the M's are linear forms. We suppose that $q>h+1$, where h is the index of Q. We may suppose that Q is written as

$$\sum_{i=1}^{h} x_i^2 - \sum_{i=h+1}^{r} x_i^2.$$

We set $x_1= \cdots =x_h=0$. Since $M_1, \cdots, M_q$ are linearly independent to begin with, at most h of these vanish. We may assume, therefore, that M_{h+1} and M_{h+2} are linearly independent of $x_1, \cdots, x_h$. We also set $L=M_{q+1}= \cdots =M_r=0$. We have

$$(5.1) \qquad - \sum_{i=h+1}^{r} x_i^2 \equiv \sum_{i=h+1}^{q} M_i^2.$$

We have imposed at most $r-q+h+1$ independent conditions on the variables in Q. It follows that the left and right members in (5.1) do not vanish identically, whence we have a contradiction.

The theorems above and well known rank theory now imply that the rank, characteristic, and, if the field K is real, also the index of a quadratic form Q, have the common property that they change at most by 1 under addition of a term λL^2, L linear, to Q.

Bibliography

1. Ernst Witt, *Theorie der quadratischen Formen in beliebigen Körpern*, J. Reine Angew. Math. vol. 176 (1937) pp. 31–44.

2. Alfred Loewy, *Ueber Scharen reeller quadratischer und Hermitescher Formen*, ibid. vol. 122 (1900) pp. 53–72.

3. Rufus Oldenburger, *Expansions of quadratic forms*, Bull. Amer. Math. Soc. vol. 49 (1942) pp. 136–141.

4. Leonard E. Dickson, *On quadratic forms in a general field*, ibid. vol. 14 (1907) pp. 108–115.

5. Helmut Hasse, *Über die Darstellbarkeit von Zahlen durch quadratischen Formen im Körper der rationalen Zahlen*, J. Reine Angew. Math. vol. 152 (1923) pp. 129–148.

Illinois Institute of Technology,
Chicago, Ill.

ON THE OSCILLATION OF DIFFERENTIAL TRANSFORMS. IV
JACOBI POLYNOMIALS[1]

BY

G. SZEGÖ

1. Introduction. In a paper in the Trans. Amer. Math. Soc.[2], E. Hille proved the following

THEOREM A. *Let* $\alpha \geqq 0$, $\beta \geqq 0$, $c \geqq 0$. *The differential operation*

$$(1.1)\qquad \vartheta - c = (1 - x^2)D^2 + [\beta - \alpha - (\alpha + \beta + 2)x]D - c, \qquad D = d/dx,$$

does not diminish the number of the sign changes in the interval $-1 < x < +1$.

More exactly, let $y = y(x)$ be a real-valued non-constant function of x, $-1 \leqq x \leqq +1$, with a continuous second derivative (with one-sided derivatives at the end points ± 1). Then the number of the sign changes of $Y = (\vartheta - c)y$ in -1, $+1$ is not less than that of y in the same interval[3].

First let us observe that under the conditions mentioned Y cannot vanish identically—this being true even for $\alpha > -1$, $\beta > -1$. More precisely, the solutions of the differential equation $(\vartheta - c)y = 0$ which are not identically zero cannot have a continuous second derivative in the closed interval $-1 \leqq x \leqq +1$, provided $c > 0$; in the case $c = 0$ the solution $y = \text{const.}$ is the only one of the kind mentioned[4]. Indeed, let us assume that $c > 0$, and let $u(x)$ and $v(x)$ be the solutions of the differential equation mentioned regular at $x = +1$ and $x = -1$, respectively, and satisfying the condition $u(+1) = v(-1) = 1$ [see (2.1)]. Then by means of the table in §2 below we conclude that $u(x)$ and $v(x)$ are linearly independent [$u'(x) \to \infty$, $v'(x) = O(1)$ as $x \to -1+0$ and $u'(x) = O(1)$, $v'(x) \to \infty$ as $x \to 1-0$]. Moreover $\{c_1u(x) + c_2v(x)\}' \to \infty$ either for $x \to -1+0$ or for $x \to 1-0$ (or in both cases) unless $c_1 = c_2 = 0$.

In the same paper E. Hille proved by means of Theorem A the special case $c = 0$ of the following

THEOREM B. *Let* $\alpha \geqq 0$, $\beta \geqq 0$, $c \geqq 0$ *and let* ϑ *have the same meaning as in*

Presented to the Society, February 27, 1943; received by the editors October 17, 1942.

(1) See the previous papers of this series by G. Szegö, E. Hille and A. C. Schaeffer, in the Trans. Amer. Math. Soc. vols. 52, 53 (1942–1943). (Cf. below, loc. cit. footnotes 2 and 6.)

(2) E. Hille, *On the oscillation of differential transforms.* II. *Characteristic series of boundary value problems*, Trans. Amer. Math. Soc. vol. 52 (1942) pp. 463–497; see §2.8.

(3) Regarding the definition of the number of sign changes see, G. Pólya and G. Szegö, *Aufgaben und Lehrsätze aus der Analysis*, vol. 2, 1925, p. 40.

(4) G. Szegö, *Orthogonal polynomials*, Amer. Math. Soc. Colloquium Publications, vol. 23, 1939; see p. 61, (4.2.6).

Theorem A. *We denote by $f(x)$ a real-valued function possessing derivatives of all orders in $-1 \leqq x \leqq +1$. If the number of the sign changes of the functions $(\vartheta - c)^k f(x)$, $k = 1, 2, 3, \cdots$, is bounded, say at most N, then $f(x)$ is a polynomial of degree at most N.*

The purpose of the present note is to prove

THEOREM A′. *Theorem* A *remains true under the more general condition $\alpha > -1$, $\beta > -1$, $c \geqq 0$.*

THEOREM B′. *Let α and β be arbitrary real, $c \geqq 0$. If $f(x)$ satisfies the conditions of Theorem* B, *$f(x)$ must be a polynomial of degree at most $N + \gamma$. Here the constant $\gamma = \gamma(\alpha, \beta, c)$ depends only on α, β and c.*

Assuming $\alpha > -1$, $\beta > -1$, Theorem B′ (with $\gamma = 0$) can be derived from Theorem A′ in a manner used first by G. Pólya and N. Wiener in case of Fourier series[5] and applied later to numerous other instances by E. Hille (loc. cit.). We prefer however a direct proof of Theorem B′ based on an idea which was used in the first paper of the present series[6].

2. **Proof of Theorem** A′. First we assume $c > 0$. Let $u(x)$ be the uniquely determined solution of $(\vartheta - c)y = 0$ which is regular at $x = +1$ and for which $u(+1) = 1$ holds; we have as well known

$$\begin{aligned} u(x) &= F(k, k'; l; (1-x)/2) \\ &= \sum_{n=0}^{\infty} \frac{k(k+1) \cdots (k+n-1)k'(k'+1) \cdots (k'+n-1)}{l(l+1) \cdots (l+n-1) \cdot 1 \cdot 2 \cdots n} ((1-x)/2)^n \end{aligned} \tag{2.1}$$

where k and k' are the roots of the quadratic equation $k(-k+\alpha+\beta+1) = c$ and $l = \alpha + 1$. Since $(k+\nu)(k'+\nu) = \nu(\nu+\alpha+\beta+1) + c > 0$, $\nu = 0, 1, 2, \cdots$, we have $u(x) > 0$ and $u'(x) < 0$ in $-1 < x \leqq +1$. Incidentally, k and k' are different from $0, -1, -2, \cdots$; $l > 0$.

Let us investigate the behavior of $u(x)$ and $u'(x)$ as $x \to -1 + 0$. Since

$$\begin{aligned} \frac{k(k+1) \cdots (k+n-1)k'(k'+1) \cdots (k'+n-1)}{l(l+1) \cdots (l+n-1) \cdot 1 \cdot 2 \cdots n} &\cong \frac{\Gamma(l)}{\Gamma(k)\Gamma(k')} n^{k+k'-l-1} \\ &= \frac{\Gamma(l)}{\Gamma(k)\Gamma(k')} n^{\beta-1}, \quad n \to \infty, \end{aligned} \tag{2.2}$$

Cesàro's theorem[7] can be applied to $u(x)$ provided $\beta \geqq 0$ and to $u'(x)$ provided $\beta > -1$. We obtain the following table:

---

(5) G. Pólya and N. Wiener, *On the oscillation of the derivatives of a periodic function*, Trans. Amer. Math. Soc. vol. 52 (1942) pp. 249–256.

(6) G. Szegö, *On the oscillation of differential transforms*. I, Trans. Amer. Math. Soc. vol. 52 (1942) pp. 450–462.

(7) See, for instance, G. Pólya and G. Szegö, *Aufgaben und Lehrsätze aus der Analysis*, vol. 1, 1925, p. 14, Problem 85.

(2.3)

	$u(x)\sim$	$-u'(x)\sim$
$\beta>0$	$(1+x)^{-\beta}$	$(1+x)^{-\beta-1}$
$\beta=0$	$-\log(1+x)$	$(1+x)^{-1}$
$-1<\beta<0$	1	$(1+x)^{-\beta-1}$

The symbol $f(x)\sim g(x)$ means that $f(x)/g(x)$ approaches a positive limit as $x\to-1+0$.

We also note the identity

$$(2.4)\qquad \begin{cases} Y = (\vartheta - c)y = (1-x)^{-\alpha}(1+x)^{-\beta}\{u(x)\}^{-1}t'(x), \\ t(x) = H(x)(y'u - yu'), \qquad H(x) = (1-x)^{\alpha+1}(1+x)^{\beta+1}. \end{cases}$$

Now let y have N sign changes in $-1<x<+1$, $N>0$; then N abscissae α_ν exist, $\alpha_0=1>\alpha_1>\alpha_2>\cdots>\alpha_N>\alpha_{N+1}=-1$, such that y is alternately less than or equal to 0 and greater than or equal to 0 in the intervals $\alpha_{\nu+1}$, α_ν without being identically zero in these intervals. We may assume that in an arbitrary small left-hand neighborhood of α_ν there are abscissae for which $y\neq 0$, $1\leqq\nu\leqq N$. (By this condition the α_ν are uniquely determined.) Obviously $y(\alpha_\nu)=0$, $1\leqq\nu\leqq N$. Then by Rolle's theorem we conclude the existence of at least $N-1$ zeros for $u^2(y/u)'=y'u-yu'$ hence also for $t(x)$ between α_1 and α_N separating the abscissae α_ν; in addition $\lim t(x)=0$ as $x\to1-0$.

But $t(x)$ must have also a zero in $-1<x<\alpha_N$. Assume the contrary, for instance $t(x)<0$ or $(y/u)'<0$ in $-1<x<\alpha_N$. Then y/u is decreasing in this interval and since $y(\alpha_N)=0$ we must have $y>0$ in $-1<x<\alpha_N$ and $y>hu$ in $-1<x\leqq\alpha_N-\epsilon\,[0<\epsilon<\alpha_N+1,\ h=h(\epsilon)>0]$.

In case $\beta\geqq0$ we conclude that $y\to+\infty$ as $x\to-1+0$ [see table (2.3)] which is a contradiction.

In case $-1<\beta<0$ we obtain $y>h'$ $(h'>0)$ for $-1<x\leqq\alpha_N-\epsilon$. But in this case $-u/u'\sim(1+x)^{\beta+1}\to0$ as $x\to-1+0$ so that

$$(2.5)\qquad \frac{t(x)}{-(1+x)^{\beta+1}yu'} = \frac{(1-x)^{\alpha+1}(1+x)^{\beta+1}(y'u-yu')}{-(1+x)^{\beta+1}yu'} \to 2^{\alpha+1},$$

hence $t(x)>0$ when x is sufficiently near -1. This is again a contradiction.

Recapitulating, we have found certain zeros $\beta_0, \beta_1, \cdots, \beta_N$ of $t(x)$ satisfying the inequalities $\beta_0=1>\beta_1>\cdots>\beta_{N-1}>\beta_N>-1$ and $\alpha_{\nu+1}<\beta_\nu<\alpha_\nu$, $1\leqq\nu\leqq N$. Repeated application of Rolle's theorem furnishes at least N sign changes of Y. Note that $t(x)$ cannot be identically 0 in $\beta_{\nu+1}$, β_ν since this would imply $y/u\equiv$const., hence $y\equiv0$ on account of $y(\alpha_{\nu+1})=0$. But $y\neq0$ at suitable points to the left from $\alpha_{\nu+1}$.

The remaining case $c=0$ can easily be settled. The identity (2.4) holds then with $u(x)=1$, that is, $t(x)=H(x)y'$. In this case $t(x)$ has at least $N-1$ zeros in the interior of -1, $+1$ and in addition the zeros $x=\pm1$.

3. **Proof of Theorem** B′. Let us start with certain preliminary remarks on Jacobi polynomials $P_n^{(\alpha,\beta)}(x)$. For arbitrary real values of α and β we use the definition [see Szegö, loc. cit.(4) p. 61, (4.21.2)]

$$
(3.1)\quad \begin{cases} P_0^{(\alpha,\beta)}(x)=1; \\ P_n^{(\alpha,\beta)}(x)=C_{n+\alpha,n}F(-n,\ n+\alpha+\beta+1;\ \alpha+1;\ (1-x)/2) \\ \qquad =(n!)^{-1}\sum_{\nu=0}^{n} C_{n,\nu}(n+\alpha+\beta+1)\cdots(n+\alpha+\beta+\nu)(\alpha+\nu+1) \\ \qquad\qquad \cdots(\alpha+n)((x-1)/2)^{\nu}, \qquad n\geqq 1. \end{cases}
$$

Then $y=P_n^{(\alpha,\beta)}(x)$ satisfies the differential equation $(\vartheta+n(n+\alpha+\beta+1))y=0$ [Szegö, loc. cit. p. 59, (4.2.1)]. Furthermore, except for an additive constant [loc. cit. p. 62, (4.21.7)]

$$
(3.2)\qquad \int P_n^{(\alpha,\beta)}(x)dx = 2(n+\alpha+\beta)^{-1}P_{n+1}^{(\alpha-1,\beta-1)}(x).
$$

We also note Rodrigues' formula [loc. cit. p. 66, (4.3.1)]

$$
(3.3)\qquad \begin{aligned} &(1-x)^{\alpha}(1+x)^{\beta}P_n^{(\alpha,\beta)}(x) \\ &\qquad = (-1)^n(2^n n!)^{-1}(d/dx)^n\{(1-x)^{n+\alpha}(1+x)^{n+\beta}\}. \end{aligned}
$$

From (3.1) we see that $P_n^{(\alpha,\beta)}(x)$, $n\geqq 1$, is of the precise degree n provided $\alpha+\beta\neq -2, -3, -4, \cdots$. If $\alpha+\beta=-l-1$, l positive integer, $P_n^{(\alpha,\beta)}(x)$ is still of the precise degree n provided $n>l$.

In case $\alpha>-1$, $\beta>-1$ we conclude from (3.3) in the familiar manner the orthogonality relation

$$
(3.4)\qquad \int_{-1}^{+1}(1-x)^{\alpha}(1+x)^{\beta}P_n^{(\alpha,\beta)}(x)q(x)dx = 0
$$

where $q(x)$ is an arbitrary polynomial of degree $n-1$. Now let α and β be arbitrary real and let m be the smallest non-negative integer such that $\alpha+m>-1$, $\beta+m>-1$. Taking $n\geqq 2m+1$ and $q(x)=(1-x^2)^m r(x)$ where $r(x)$ is an arbitrary polynomial of degree $n-2m-1$ we find that for this particular type of polynomials $q(x)$ the orthogonality relation (3.4) still holds.

Under the same condition we have [loc. cit. p. 62, (4.21.6), p. 67, (4.3.3)]

$$
(3.5)\qquad \begin{aligned} &\int_{-1}^{+1}(1-x)^{\alpha+m}(1+x)^{\beta+m}P_n^{(\alpha,\beta)}(x)x^{n-2m}dx \\ &\qquad = (-1)^m 2^{n+\alpha+\beta+1}\frac{\Gamma(n+\alpha+1)\Gamma(n+\beta+1)}{\Gamma(2n+\alpha+\beta+2)} \neq 0. \end{aligned}
$$

After these preliminaries we proceed to the proof of Theorem B′. First let us exclude the case $\alpha+\beta=-l-1$, l positive integer. We expand $f^{(m)}(x)$ in a series of Jacobi polynomials $P_n^{(\alpha+m,\beta+m)}(x)$:

$$f^{(m)}(x) = \sum_{n=0}^{\infty} f_n P_n^{(\alpha+m,\beta+m)}(x). \tag{3.6}$$

Term-by-term integration and use of (3.2) furnishes

$$\begin{aligned} f(x) = \phi(x) + \sum_{n=0}^{\infty} 2^m \{ & (n+\alpha+\beta+2m)(n+\alpha+\beta+2m-1) \\ & \cdots (n+\alpha+\beta+m+1)\}^{-1} f_n P_{n+m}^{(\alpha,\beta)}(x) \end{aligned} \tag{3.7}$$

where $\phi(x)$ is a polynomial of degree $m-1$ [for $m=0$ we have $\phi(x)=0$]. Since in this case $P_n^{(\alpha,\beta)}(x)$ is of the precise degree n we can write

$$f(x) = \sum_{n=0}^{\infty} \phi_n P_n^{(\alpha,\beta)}(x). \tag{3.8}$$

Obviously

$$(\vartheta - c)^k f(x) = \sum_{n=0}^{\infty} (-1)^k [c + n(n+\alpha+\beta+1)]^k \phi_n P_n^{(\alpha,\beta)}(x). \tag{3.9}$$

Now let k belong to a certain infinite sequence such that the corresponding functions $(\vartheta-c)^k f(x)$ have a fixed number, M say, sign changes; $M \leqq N$(8). We denote the abscissae at which these sign changes take place by $x_1, x_2, \cdots, x_M$; $x_\nu = x_\nu(k)$. Then if $\delta=+1$ or -1 is properly chosen,

$$\begin{aligned} \delta \int_{-1}^{+1} (1-x)^{\alpha+m}(1+x)^{\beta+m} & \{(\vartheta-c)^k f(x)\}(x-x_1) \\ & \cdots (x-x_M)(1 \pm x^\rho)dx > 0. \end{aligned} \tag{3.10}$$

Here ρ is an arbitrary non-negative integer and δ does not depend on ρ. Substituting for $(\vartheta-c)^k f(x)$ its expansion (3.9) the arising integrals will all vanish provided $n > 2m+M+\rho$. However for $n=n'=2m+M+\rho$ we obtain

$$\pm\, \delta(-1)^k [c + n'(n'+\alpha+\beta+1)]^k \cdot \phi_{n'} \int_{-1}^{+1} (1-x)^{\alpha+m}(1+x)^{\beta+m} P_{n'}^{(\alpha,\beta)}(x) x^{M+\rho} dx,$$

and the last integral is different from 0 because of (3.5). Hence if $\phi_{n'} \neq 0$ we find for $k \to \infty$

$$[c + n'(n'+\alpha+\beta+1)]^k = O(1) \max_{0 \leq \nu \leq n'-1} |c + \nu(\nu+\alpha+\beta+1)|^k$$

(8) From here on we use the argument of the paper cited in footnote 6.

which is impossible provided

$$\left| c + n'(n' + \alpha + \beta + 1) \right| > \max_{0 \leq \nu \leq n'-1} \left| c + \nu(\nu + \alpha + \beta + 1) \right|.$$

This is the case if $n' \geqq n_0 = n_0(\alpha, \beta, c)$.

The previous argument furnishes $\phi_n = 0$ for $n \geqq 2m + M$, $n \geqq n_0$, which is equivalent to the assertion of Theorem B′.

In case $\alpha + \beta = -l - 1$, l positive integer, this proof needs a slight modification. We integrate then only the terms $n \geqq m+1$ in (3.6) and conclude (3.7) with the modification that the summation is now extended over the range $n \geqq m+1$ and $\phi(x)$ is a polynomial of degree $2m$. [The expression in the braces of (3.7) is then positive since $2m + \alpha + \beta + 2 > 0$.] As a further addition to the previous argument we have to show that

$$(\vartheta - c)^k \phi(x) = O(1) \left| c + n'(n' + \alpha + \beta + 1) \right|^k, \qquad k \to \infty,$$

uniformly for $-1 \leqq x \leqq +1$ provided n' is sufficiently large, $n' \geqq n_1 = n_1(\alpha, \beta, c)$. But $(\vartheta - c)^k \phi(x)$ is a polynomial of degree $2m$ and the last assertion follows if we can show that the coefficients of this polynomial have moduli at most RS^k; here $R > 0$ depends on $f(x)$, α, β, c and $S > 0$ depends only on α, β, c. Now

$$(3.11) \qquad \begin{aligned} (\vartheta - c)x^h = {} & h(h-1)(1-x^2)x^{h-2} \\ & + h[\beta - \alpha - (\alpha + \beta + 2)x]x^{h-1} - cx^h; \end{aligned}$$

hence with arbitrary constants λ_h

$$(3.12) \qquad (\vartheta - c) \sum_{h=0}^{2m} \lambda_h x^h \ll S \cdot \max \left| \lambda_h \right| \cdot \sum_{h=0}^{2m} x^h$$

where

$$(3.13) \qquad S = 2 \cdot 2m(2m-1) + 2m \left| \beta - \alpha \right| + 2m \left| \alpha + \beta + 2 \right| + \left| c \right|.$$

This furnishes the statement by taking for R the maximum modulus of the coefficients of $\phi(x)$ and choosing S according to (3.13). ∎

Theorems B and B′ remain of course true if the condition regarding $(\vartheta - c)^k f(x)$ is satisfied only for an infinite number of values of k.

STANFORD UNIVERSITY,
STANFORD UNIVERSITY, CALIF.

ON STRUCTURES OF INFINITE MODULES

BY

R. E. JOHNSON

Much of the literature on the structures of modules applies to those modules which possess a finite basis. The present paper is the development of a structure theory for particular infinite modules with countable bases. Generality of results is not as much the aim of the paper as is the application to problems concerning infinite matrices.

For a commutative field P, Ξ is assumed to be a universal P-module which has a countable P-basis. A principal ideal ring Q which contains P is considered as an operator domain of Ξ. Then the main topic studied is under what conditions submodules of Ξ have proper Q-bases.

In the first place, a complete characterization is given for the proper Q-bases of any Q-submodule of Ξ. This is represented as an infinite matrix, and is called the characteristic matrix of the submodule.

The finite case is studied in the third section. The results obtained are comparable with those of Ingraham and Wolf [3](1) and Chevalley [1]. The principal theorem is that every Q-module which possesses a finite Q-basis has a proper Q-basis.

The concepts of primitivity—defined somewhat as Chevalley defines it—and index play an important role in determining conditions for a Q-module to have a proper Q-basis. In order to find these conditions, the non-regular elements H of Ξ are split from Ξ. The resulting Q-module Ξ/H is regular. Then necessary and sufficient conditions are found for both H and Ξ/H to have a proper Q-basis.

If the operator domain of Ξ be considered as $Q/(m)$, m not a unit of Q, then in the fifth section it is seen that Ξ possesses a proper $Q/(m)$-basis.

As an application of these results, Ξ is taken to be the set of all vectors over P of order type ω which are finitely nonzero. The total operator domain of Ξ is a certain ring of infinite matrices, $\mathfrak{M}_\omega$. Then any element A of $\mathfrak{M}_\omega$ can be transformed into a direct sum of finite matrices only if Ξ has a proper $P[A]$-basis.

The algebraic theory assumed herein can be found in almost any book on modern algebra—specific attention is called to MacDuffee [4] and Zassenhaus [5].

1. **Introduction.** Let P denote a commutative field, and Q a principal ideal

Presented to the Society, April 18, 1942; received by the editors April 28, 1942, and, in revised form, October 13, 1942.

(1) The numbers in brackets refer to the bibliography at the end of the paper.

ring which contains P[3]. The universal P-module (linear set over P) of all modules discussed below will be labelled by Ξ. It is assumed to have Q as an operator domain. While Ξ will in general have an infinite number of elements, yet only finite sums are ever considered. Small Greek letters will always denote elements of Ξ, capital Greek letters will stand for subsets of Ξ, and small Latin letters will be used for elements of Q.

For subsets $\Xi_1, \Xi_2, \cdots$ of Ξ, $\Xi_1 \vee \Xi_2 \vee \cdots$ is used to denote the least P-module in Ξ which contains all Ξ_i, while $\Xi_1 \wedge \Xi_2 \wedge \cdots$ denotes the set-theoretic intersection of all Ξ_i. If the Ξ_i, $i=1, 2, \cdots$, are P-modules, and $\Xi' = \Xi_1 \vee \Xi_2 \vee \cdots$, then Ξ' is the supplementary sum of the Ξ_i, written

$$\Xi' = \Xi_1 + \Xi_2 + \cdots$$

in case the representation of every element of Ξ' by sums of elements in the Ξ_i is unique. This is equivalent to the condition

$$\Xi_k \wedge (\Xi_1 \vee \Xi_2 \vee \cdots \vee \Xi_{k-1}) = 0, \qquad k = 2, 3, \cdots.$$

For $\Xi_2 \subset \Xi_1$, $\Xi_1 - \Xi_2$ is the set of all elements of Ξ_1 not in Ξ_2.

The universal P-module Ξ is said to have the P-basis $(\xi_1, \xi_2, \cdots)$, finite or infinite, if

$$\Xi = P[\xi_1, \xi_2, \cdots],$$

this last being the set of all finite combinations $\sum a_i\xi_i$, $a_i \in P$. If the set $(\xi_1, \xi_2, \cdots)$ is P-linearly independent—that is, for every finite sum $\sum a_i\xi_i = 0$, all $a_i = 0$—this basis is called regular. In this case one has

$$\Xi = P\xi_1 + P\xi_2 + \cdots, \quad P\xi_i = P[\xi_i].$$

The following axiom is assumed throughout the paper.

FUNDAMENTAL AXIOM. *The universal module Ξ has a countable P-basis.*

A consequence of this axiom (see Ingraham [2]) is

THEOREM 1.1. *Every P-submodule Ξ_1 of Ξ has a proper P-basis.*

The set $(\xi_1, \xi_2, \cdots)$ is a Q-basis for Ξ_1 in case $\Xi_1 = Q[\xi_1, \xi_2, \cdots]$. The set Ξ_1 is a Q-module in case $Q\xi \subset \Xi_1$ for every $\xi \in \Xi_1$.

DEFINITION 1.2. *A set of elements $(\xi_1, \xi_2, \cdots)$ is a proper Q-basis of Ξ_1, and is thus Q-linearly independent, if and only if*

(1) $$\xi_i \neq 0, \qquad i = 1, 2, \cdots,$$

and

(2) $$\Xi_1 = Q\xi_1 + Q\xi_2 + \cdots.$$

(3) As a particular instance of these concepts, we can consider P as the rational field and Q the ring $P[x]$ of all polynomials in the indeterminate x with coefficients in P. The first part of §6 might well be read first to give one a concrete example of the sets Q and Ξ.

An element a of Q is an annihilator of ξ if $a\xi=0$. If a_1 and a_2 are two annihilators of ξ, then $b_1a_1+b_2a_2$ is an annihilator of ξ for any two elements b_1 and b_2 of Q. Thus the set of all annihilators of ξ form an ideal. Since Q is by assumption a principal ideal ring, this ideal is principal of the form (h). As h divides all annihilators of ξ, it will be called a minimum annihilator of ξ. It is unique up to a unit factor.

DEFINITION 1.3. *If $a\xi=0$ implies $a=0$, then ξ is called regular. If every nonzero element of a Q-module Ξ_1 is regular, then Ξ_1 is called regular.*

If ξ and η are two non-regular elements of Ξ with minimum annihilators a_1 and a_2, respectively, then for any elements b_1 and b_2 of Q, $b_1\xi+b_2\eta$ is annihilated by a_1a_2. Thus the set of all non-regular elements of Ξ forms a Q-module which we shall label H. Suppose $(\eta_1, \eta_2, \cdots)$ is a P-basis for H, with h_i the minimum annihilator of η_i. Then we have defined a set $(p_1, p_2, \cdots)$ of primes of Q, which are all the distinct prime factors of the h_i, $i=1, 2, \cdots$ (p and q are not distinct if $p=qc$, c a unit). Now any element η of H is of the form $\eta=\sum_{i=1}^{n}a_i\eta_i$, $a_i\in P$, so η is annihilated by some product of the primes p_j. For any element p of Q, let H_p be the set of all elements of H annihilated by some power of p. Then H_p is a Q-module.

THEOREM 1.4. $H=H_{p_1}+H_{p_2}+\cdots$.

To prove this, let η be any element of H, and let $\prod_{i=1}^{n}p_i^{t_i}$ be its minimum annihilator. Then there exist elements s_i such that

$$\sum_{j=1}^{n} s_j \prod_{i=1, i\neq j}^{n} p_i^{t_i} = 1.$$

If

$$\eta_j = s_j \prod_{i=1, i\neq j}^{n} p_i^{t_i}\eta,$$

then $\eta_i\in H_{p_i}$, $\eta=\eta_1+\eta_2+\cdots+\eta_n$.

As an immediate consequence of this theorem, we have that for any Q-module $H_1\subset H$,

$$H_1 = H_1 \wedge H_{p_1} + H_1 \wedge H_{p_2} + \cdots .$$

Any P-module Ξ_1 is fundamentally an abelian group, so that the quotient group Ξ_1/Ξ_2 is well defined for any P-submodule Ξ_2 of Ξ_1, and is itself a P-module. Likewise, if Ξ_1 and Ξ_2 are Q-modules with $\Xi_2\subset\Xi_1$, then Ξ_1/Ξ_2 is a Q-module.

If H_1 is the set of all non-regular elements of the Q-module Ξ_1, then Ξ_1/H_1 has elements of the form $(\xi+H_1)$ (this should not be confused with the supplementary sum—it means the set of all elements of the form $\xi+\eta$, $\eta\in H_1$) for $\xi\in\Xi_1$. Suppose $a(\xi+H_1)=(H_1)$ for some element $a\in Q$, $\xi\in\Xi_1-H_1$. Then $a\xi\in H_1$ so that $a=0$. Thus Ξ_1/H_1 is a regular Q-module.

2. **Invariants of proper Q-bases.** Let Ξ_1, a Q-module, have a proper Q-basis $(\eta_1, \eta_2, \cdots, \xi_1, \xi_2, \cdots)$, the ξ_i being regular elements, and η_i having minimum annihilator q_i, $i=1, 2, \cdots$. If $H_1=Q[\eta_1, \eta_2, \cdots]$, $\Omega_1=Q[\xi_1, \xi_2, \cdots]$, then $\Xi_1=H_1+\Omega_1$, $H_1\subset H$.

Definition 2.1. *The infinite matrix* $(n_{rs}; r=0, 1, 2, \cdots; s=1, 2, \cdots)$, *in which* n_{rs} *is the number of the* q_j, $j=1, 2, \cdots$, *divisible by* p_r^s *but not by* p_r^{s+1} *for* $r, s=1, 2, \cdots$, *and* n_{0s} *is the cardinal number of the set* $(\xi_1, \xi_2, \cdots)$, $s=1, 2, \cdots$, *is called the characteristic matrix*(3) *of the proper Q-basis* $(\eta_1, \eta_2, \cdots, \xi_1, \xi_2, \cdots)$. *The elements of this matrix are integers or* $\aleph_0$.

Theorem 2.2. *The characteristic matrix is an invariant of the class of all proper Q-bases of* Ξ_1.

To prove this, let $(\eta_{j1}, \eta_{j2}, \cdots, \xi_{j1}, \xi_{j2}, \cdots)$, $j=1, 2$, be two proper Q-bases of Ξ_1, with the ξ_{ji} regular and η_{ji} annihilated minimally by q_{ji}, $i=1, 2, \cdots$. The number of elements of the set $(q_{j1}, q_{j2}, \cdots)$ divisible by p_r^s but not by p_r^{s+1} is denoted by n_{jrs}, $r, s=1, 2, \cdots$. The cardinal number of the set $(\xi_{j1}, \xi_{j2}, \cdots)$ is denoted by n_{j0s}, so that $n_{j01}=n_{j02}=\cdots$. We shall first show that $n_{111}=n_{211}$.

Select r_{ij} from Q so that $q_{ij}=p_1^{t_{ij}}r_{ij}$ with the greatest common divisor of p_1 and r_{ij}, denoted (p_1, r_{ij}), equal to 1. The nonzero elements of the set $(r_{j1}\,\eta_{j1}, r_{j2}\,\eta_{j2}, \cdots)$ are Q-linearly independent: let us relabel these nonzero elements $(\alpha_{j1}, \alpha_{j2}, \cdots)$. Then

$$Q\alpha_{11}+Q\alpha_{12}+\cdots=Q\alpha_{21}+Q\alpha_{22}+\cdots, \tag{1}$$

this being the set H_{p_1} of all elements of H_1 annihilated by some power of p_1.

If $n_{111}=n_{211}=\aleph_0$, the first step of the proof is concluded. Thus assume $n_{111}=n<\aleph_0$. Let us separate the α_{ji} into the sets $(\beta_{j1}, \beta_{j2}, \cdots)$, $(\gamma_{j1}, \gamma_{j2}, \cdots)$ with the first set being all the α_{ji} annihilated by p_1, and the second set the remainder of the α_{ji}. It is observed that no member of the set $(\beta_{11}, \beta_{12}, \cdots, \beta_{1n})$ could be in the module $Q[\gamma_{21}, \gamma_{22}, \cdots]$. For if

$$\beta_{11}=\sum_{i=1}^{t} c_i\gamma_{2i},$$

then each c_i must be divisible by p_1 as p_1 annihilates β_{11}, but does not annihilate any γ_{ji}. Thus $\beta_{11}=p_1\gamma$. However, $\gamma\in Q[\beta_{11}, \beta_{12}, \cdots, \gamma_{11}, \gamma_{12}, \cdots]$, say

$$\gamma=\sum_{i=1}^{s} d_i\beta_{1i}+\sum_{i=1}^{r} e_i\gamma_{1i},$$

so that $\beta_{11}=\sum_{i=1}^{r} p_1e_i\gamma_{1i}$, which is impossible. From (1), we have

(3) This bears no relationship with the ordinary concept of characteristic matrix. The name was chosen because of the connection between this matrix and the characteristic divisors of certain infinite matrices.

$$\beta_{1i} = \sum_{j=1}^{m_i} b_{1ij}\beta_{2j} + p_1\delta_{1i}, \qquad \delta_{1i} \in Q[\gamma_{21}, \gamma_{22}, \cdots],$$

$$\beta_{2i} = \sum_{j=1}^{n} b_{2ij}\beta_{1j} + p_1\delta_{2i}, \qquad \delta_{2i} \in Q[\gamma_{11}, \gamma_{12}, \cdots].$$

A substitution yields, if we let m be the maximum m_i, and $b_{1ij}=0$, $m_i<j\leqq m$,

$$\tag{2} \begin{aligned} \beta_{1i} &= \sum_{j=1}^{m}\sum_{k=1}^{n} b_{1ij}b_{2jk}\beta_{1k} + p_1\delta_1, \\ \beta_{2i} &= \sum_{j=1}^{n}\sum_{k=1}^{m} b_{2ij}b_{1jk}\beta_{2k} + p_1\delta_2. \end{aligned}$$

From the Q-linear independence of the sets $(\beta_{j1}, \beta_{j2}, \cdots, \gamma_{j1}, \gamma_{j2}, \cdots)$ one must have $p_1\delta_1 = p_1\delta_2 = 0$. If the matrices B_1 and B_2 are defined as

$$B_1 = (b_{1rs};\ r = 1, \cdots, n;\ s = 1, \cdots, m),$$
$$B_2 = (b_{2rs};\ r = 1, \cdots, m;\ s = 1, \cdots, n),$$

then from (2) one can conclude (using I_k as the unit matrix of order k^2)

$$B_1B_2 = I_n, \qquad B_2B_1 = I_m.$$

However B_1 and B_2 have elements in the field $Q/(p_1)$, so that m and n must be equal. Also, from (2), for $t>n$,

$$\beta_{2t} = \sum_{j=1}^{n}\sum_{k=1}^{m} b_{2tj}b_{1jk}\beta_{2k}$$

which is impossible, as the set $(\beta_{21}, \beta_{22}, \cdots)$ is Q-linearly independent. The conclusion is that there are n elements in the set $(\beta_{21}, \beta_{22}, \cdots)$, so that $n_{111} = n_{211}$.

In order to show that $n_{11t} = n_{21t}$, consider the set $p_1^{t-1}H_{p_1}$, which has the two proper (if we exclude the zero elements) Q-bases $(p_1^{t-1}\alpha_{j1}, p_1^{t-1}\alpha_{j2}, \cdots)$, $j=1, 2$. From the paragraph above, the number of nonzero elements annihilated by p_1 in each basis is the same. This number must be the number of q_{ji} divisible exactly by p_1^t—so that $n_{11t} = n_{21t}$.

We need only to show that $n_{101} = n_{201}$ to complete the proof. Let

$$\xi_{1j} = \sum_k b_{1jk}\eta_{2k} + \sum_{k=1}^{m_{1j}} d_{1jk}\xi_{2k}, \qquad j = 1, 2, \cdots, n_{101},$$

$$\xi_{2j} = \sum_k b_{2jk}\eta_{1k} + \sum_{k=1}^{m_{2j}} d_{2jk}\xi_{1k}, \qquad j = 1, 2, \cdots, n_{201}.$$

Case 1. Let $n_{101} = \aleph_0$, n_{201} finite: if m be the maximum m_{2j}, $j = 1, 2, \cdots, n_{201}$, then $\Xi_1 \subset Q[\eta_{11}, \eta_{12}, \cdots, \xi_{11}, \xi_{12}, \cdots, \xi_{1m}]$ which is impossible.

Case 2. Let n_{101}, n_{201} both be finite: then there exists an element $c \in Q$ such that

$$c\xi_{1j} = c\sum_{i=1}^{n_{201}} d_{1ji}\xi_{2i}, \qquad j = 1, 2, \cdots, n_{101},$$

$$c\xi_{2j} = c\sum_{i=1}^{n_{101}} d_{2ji}\xi_{1i}, \qquad j = 1, 2, \cdots, n_{201}.$$

If $D_k = (d_{krs})$, $k = 1, 2$, then by the method above

$$cD_1D_2 = cI_{n_{101}}, \qquad cD_2D_1 = cI_{n_{201}}.$$

As Q can always be imbedded in a field, we can consider D_1 and D_2 as having elements in a field, so that $n_{101} = n_{201}$. Thus the characteristic matrices of the two bases must be equal.

Consider again the proper Q-basis for Ξ_1 given at the beginning of this section. If $q_1 = cp_{i_1}^{t_1} \cdots p_{i_n}^{t_n}$, c a unit and all $t_i \geqq 1$, then let

$$\eta_{11} = p_{i_2}^{t_2} \cdots p_{i_n}^{t_n}\eta_1, \quad \eta_{12} = p_{i_1}^{t_1}p_{i_3}^{t_3} \cdots p_{i_n}^{t_n}\eta_1, \cdots, \quad \eta_{1n} = p_{i_1}^{t_1} \cdots p_{i_{n-1}}^{t_{n-1}}\eta_1.$$

We see that $Q\eta_1 = Q\eta_{11} + Q\eta_{12} + \cdots + Q\eta_{1n}$ where each η_{1i} is minimally annihilated by a power of a prime. Now this can be done for every η_i, so it is apparent that Ξ_1 has a proper Q-basis $(\beta_1, \beta_2, \cdots, \xi_1, \xi_2, \cdots)$, with the ξ_i regular and the β_i possessing powers of primes as minimum annihilators.

With the use of this basis, the completeness of the invariant characteristic matrix (n_{rs}) will be shown by the following

THEOREM 2.3. *For any set $(r_1, r_2, \cdots)$ of elements of Q such that the number of these elements divisible by p_r^s but not by p_r^{s+1} is n_{rs}, there exists a proper Q-basis $(\alpha_1, \alpha_2, \cdots, \xi_1, \xi_2, \cdots)$ of Ξ_1, with r_j the minimum annihilator of α_j, and n_{01} the number of elements in the set $(\xi_1, \xi_2, \cdots)$.*

If $r_1 = cp_{i_1}^{t_1} \cdots p_{i_k}^{t_k}$, all $t_i > 0$ and c a unit, then let j_1 be the minimum integer such that β_{j_1} has minimum annihilator $p_{i_1}^{t_1}$, and, in general, let j_s be the minimum integer such that β_{j_s} has minimum annihilator $p_{i_s}^{t_s}$. Now define $\alpha_1 = \beta_{j_1} + \beta_{j_2} + \cdots + \beta_{j_k}$, so that α_1 has minimum annihilator r_1. Discard $\beta_{j_1}, \beta_{j_2}, \cdots, \beta_{j_k}$ from the set $(\beta_1, \beta_2, \cdots)$, and with the remaining set carry through a similar process for r_2, obtaining α_2. From Theorem 2.2, there will be precisely enough elements in the set $(\beta_1, \beta_2, \cdots)$ to carry this process to completion, using all the r_i's. We can use the same regular elements in the new basis as in the old one. This completes the proof.

COROLLARY 2.4. *If all the elements of H_p are annihilated by some power of the prime p, and H_p has a proper Q-basis, then the number of elements in any proper Q-basis of H_p is an invariant of H_p.*

If a proper Q-basis has an infinite number of elements, obviously any Q-basis has an infinite number of elements. Suppose $(\eta_1, \cdots, \eta_n)$ is a proper Q-basis of H_p, and $(\alpha_1, \cdots, \alpha_m)$ any Q-basis. Then

$$\eta_i = \sum_{j=1}^{m} a_{ij}\alpha_j, \qquad \alpha_j = \sum_{k=1}^{n} b_{jk}\eta_k,$$

so that

$$\sum_{j=1}^{m} a_{ij}b_{jk} \equiv 1 \bmod p \quad \text{if} \quad j = k,$$
$$\equiv 0 \bmod p \quad \text{otherwise},$$

which is possible only if $m \geqq n$. This leads to

THEOREM 2.5. *If H_p has a proper Q-basis, then the number of elements in any Q-basis cannot be less than the number of elements in a proper Q-basis.*

3. **The finite case.** Before the general case is studied, it is necessary to consider those submodules of Ξ which have finite Q-bases.

LEMMA 3.1. *If Ξ_1 has a finite Q-basis, and $\Xi_2 \subset \Xi_1$, Ξ_2 a Q-module, then Ξ_2 has a finite Q-basis. The Q-basis for Ξ_2 can be chosen so as not to have more elements than the given Q-basis for Ξ_1.*

To prove this, let $(\xi_1, \xi_2, \cdots, \xi_n)$ be a Q-basis of Ξ_1. For any $m \leqq n$, there exists a maximum ideal $\mathfrak{a}_m \subset Q$ such that

$$\mathfrak{a}_m\xi_m \equiv 0 \bmod \Xi_2 \vee Q[\xi_1, \xi_2, \cdots, \xi_{m-1}].$$

As Q is a principal ideal ring, $\mathfrak{a}_m = (s_m)$. Select $\eta_m \in \Xi_2$ so that

$$\eta_m \equiv s_m\xi_m \bmod Q[\xi_1, \xi_2, \cdots, \xi_{m-1}], \qquad \eta_m = 0 \quad \text{if} \quad s_m = 0.$$

Suppose $\eta \in \Xi_2$, so that $\eta = \sum_{i=1}^{k} a_i\xi_i$ with $a_k\xi_k \neq 0$. Then

$$a_k\xi_k \equiv \eta \bmod Q[\xi_1, \xi_2, \cdots, \xi_{k-1}],$$

which implies

$$\eta \equiv b_k\eta_k \bmod Q[\xi_1, \xi_2, \cdots, \xi_{k-1}].$$

Thus $\eta - b_k\eta_k \in Q[\xi_1, \xi_2, \cdots, \xi_{k-1}]$, and by induction $\eta = \sum_{i=1}^{k} b_i\eta_i$. This shows that $(\eta_1, \eta_2, \cdots, \eta_n)$ is a Q-basis for Ξ_2, and establishes the lemma.

If $(\xi_1, \xi_2, \cdots, \xi_n)$ is a P-basis for the Q-module Ξ_1, then for any element $\xi \in \Xi_1$ and any element $b \in Q$, b not a unit, the set $(\xi, b\xi, \cdots, b^n\xi)$ must be P-linearly dependent. Thus $\sum_{i=1}^{n} a_ib^i\xi = 0$ for some elements $a_i \in P$; we have

LEMMA 3.2. *If Ξ_1 is a Q-module with a finite P-basis, then $\Xi_1 \subset H$.*

LEMMA 3.3. *If Ξ_1 is a Q-module with a finite Q-basis, and $\Xi_1 \subset H$, then Ξ_1 has a proper Q-basis.*

To prove this, let $\Xi_1=Q[\xi_1, \xi_2, \cdots, \xi_n]$ with ξ_i annihilated minimally by r_i, so that $\prod_{i=1}^n r_i$ annihilates Ξ_1. Let $(p_1, p_2, \cdots, p_m)$ be all the distinct prime factors of the r_i, $i=1, 2, \cdots, n$. If H_i denotes the set of all elements of Ξ_1 annihilated by some power of p_i, then from Theorem 1.4,

$$\Xi_1 = H_1 + H_2 + \cdots + H_m.$$

We will prove that each H_i has a proper Q-basis, which will imply a proper Q-basis for Ξ_1.

Denote by t_1 the minimum integer such that $p_1^{t_1}H_1=0$. Then there exists an element $\eta_1\in H_1$ for which $p_1^{t_1}$ is the minimum annihilator. Recursively, if t_k is the minimum integer such that $p_1^{t_k}H_1\equiv 0 \mod Q[\eta_1, \eta_2, \cdots, \eta_{k-1}]$, then there exists an element $\eta_k\in H_1$ which has $p_1^{t_k}$ as minimum annihilator mod $Q[\eta_1, \eta_2, \cdots, \eta_{k-1}]$. It follows that $t_1\geqq t_2\geqq \cdots \geqq t_k$.

Assume that the set $(\eta_1, \eta_2, \cdots, \eta_{k-1})$ is Q-linearly independent, and that $p_1^{t_k}\eta_k\neq 0$. From above, we must have

$$p_1^{t_k}\eta_k = \sum_{i=1}^{k-1} a_i\eta_i.$$

If this equation be multiplied by the $(t_{k-1}-t_k)$th power of p_1, it is apparent that $p_1^{t_k}|a_{k-1}$ ($p_1^{t_k}$ divides a_{k-1}). Similarly, it can be verified that $p_1^{t_k}|a_i$, $i=1, 2, \cdots, k-1$, so that $a_i=b_ip_1^{t_k}$. Then if

$$\bar{\eta}_k = \eta_k - \sum_{i=1}^{k-1} b_i\eta_i,$$

the set $(\eta_1, \eta_2, \cdots, \eta_{k-1}, \bar{\eta}_k)$ is Q-linearly independent. If this process were not finite, we would have a submodule of Ξ_1 containing an infinite proper Q-basis. This is not possible in view of Theorem 2.5 and Lemma 3.1.

THEOREM 3.4. *If Ξ_1 is a Q-module with a finite Q-basis, then Ξ_1 has a proper Q-basis.*

This follows from Lemma 3.3 if all the elements of Ξ_1 are non-regular. Thus let $H_1=\Xi_1\wedge H\neq 0$, so that Ξ_1/H_1 is regular. As Ξ_1 has a finite Q-basis, so must Ξ_1/H_1: denote this basis by $(\xi_1, \xi_2, \cdots, \xi_n)$, all $\xi_i\neq 0$.

Assume every Q-module contained in Ξ_1/H_1 which is generated by k or fewer elements has a proper (finite) Q-basis. Also assume that the set $(\xi_1, \xi_2, \cdots, \xi_k)$ is Q-linearly independent, and that ξ_{k+1} has a nonzero minimum annihilator q_{k+1} mod $Q[\xi_1, \xi_2, \cdots, \xi_k]$. Then there exist elements $q_1, \cdots, q_k\in Q$ such that

$$\sum_{i=1}^{k+1} q_i\xi_i = 0, \qquad (q_1, q_2, \cdots, q_{k+1}) = 1.$$

In the case under consideration, it is well known (see MacDuffee [4,

p. 227]) that there exists a matrix $C=(c_{rs})$ of $(k+1)^2$ elements with $c_{k+1j}=q_j$, $j=1, 2, \cdots, k+1$, such that $|C|=1$ and C has a unique inverse $B=(b_{rs})$. Let

$$\alpha_i = \sum_{j=1}^{k+1} c_{ij}\xi_j, \qquad i = 1, 2, \cdots, k.$$

Remembering that $\sum_{i=1}^{k+1} c_{k+1i}\xi_i=0$, we see that

$$\sum_{j=1}^{k} b_{ij}\alpha_j = \sum_{s,t=1}^{k+1} b_{is}c_{st}\xi_t = \xi_i, \qquad i = 1, 2, \cdots, k+1.$$

Thus $Q[\alpha_1, \alpha_2, \cdots, \alpha_k]=Q[\xi_1, \xi_2, \cdots, \xi_{k+1}]$. By assumption, this set has a proper Q-basis. An obvious induction leads to a proper Q-basis for Ξ_1/H_1. If this basis is $(\beta_1+H_1, \beta_2+H_1, \cdots, \beta_t+H_1)$, then

$$\Xi_1 = H_1 + Q\beta_1 + Q\beta_2 + \cdots + Q\beta_t,$$

and the theorem follows.

4. **The general case.** We now turn to the consideration of any Q-submodule of Ξ, and develop conditions under which it possesses a proper Q-basis.

DEFINITION 4.1. *For the Q-modules Ξ_1, Ξ_2 with $\Xi_2 \subset \Xi_1$, Ξ_2 is called primitive in Ξ_1 if, for every $\xi \in \Xi_1$ such that $p\xi \in \Xi_2$, $p\xi \neq 0$, there exists an element $\xi_1 \in \Xi_1$ for which $p\xi_1=0$, and $\xi+\xi_1 \in \Xi_2$.*

DEFINITION 4.2. *If Ξ_1 is a Q-module, then the index of an element $\xi \in \Xi_1$, written $i(\xi, \Xi_1)$, is defined as follows:*

(1) *For $\xi \in \Xi_1 \wedge H$, ξ minimally annihilated by $\prod_{j=1}^{k} p_j^{t_j}$, p_j a prime of Q and $t_j \geqq 1$ for $j=1, 2, \cdots, k$, $i(\xi, \Xi_1)$ is the maximum integer s for which there exists an element $\eta \in \Xi_1$ such that*

$$\prod_{j=1}^{k} p_j^{t_j-1}\xi = \prod_{j=1}^{k} p_j^{s_j+t_j-1}\eta, \qquad \sum_{j=1}^{k} s_j = s.$$

(2) *For ξ regular, $i(\xi, \Xi_1)$ is the maximum integer s for which there exist primes $q_1, q_2, \cdots, q_n$ in Q and $\eta \in \Xi_1$ such that*

$$\xi \equiv \prod_{j=1}^{k} q_j^{s_j}\eta \bmod \Xi_1 \wedge H, \qquad \sum_{j=1}^{k} s_j = s.$$

If in either case this maximum does not exist, $i(\xi, \Xi_1)=\infty$: $i(0, \Xi_1)=0$.

If Ξ_1 has non-regular component H_1, and $\eta \in H_1$, then obviously $i(\eta, H_1) = i(\eta, \Xi_1)$. Also, for $\Xi_2 \subset \Xi_1$, $\eta \in \Xi_2$, $i(\eta, \Xi_2) \leqq i(\eta, \Xi_1)$.

THEOREM 4.3. *Let Ξ_1 be a Q-module, and $\Xi_1 \wedge H = H_1 + H_2 + \cdots$, $H_j = \Xi_1 \wedge H_{p_j}$. Then for any $\eta \in \Xi_1 \wedge H$ so that $\eta = \eta_1 + \eta_2 + \cdots + \eta_n$, $\eta_j \in H_j$,*

$$i(\eta, \Xi_1) = \sum_{j=1}^{n} i(\eta_j, \Xi_1).$$

To prove this, we first refer back to Theorem 1.4. If η has minimum annihilator $\prod_{i=1}^{n} p_i^{t_i}$, then there exist elements $a_j \in Q$ such that

$$(1) \qquad \sum_{j=1}^{n} a_j \prod_{i=1, i\neq j}^{n} p_i^{t_i} = 1, \qquad \eta_j = a_j \prod_{i=1, i\neq j}^{n} p_i^{t_i} \eta.$$

If $i(\eta, \Xi_1) = s < \infty$, then there exists an $\alpha \in \Xi_1$ such that

$$(2) \qquad \prod_{i=1}^{n} p_i^{t_i - 1} \eta = \prod_{i=1}^{n} p_i^{t_i + s_i - 1} \alpha, \qquad \sum_{j=1}^{n} s_j = s.$$

From (1) and (2) we derive

$$p_j^{t_j - 1} \eta_j = p_j^{t_j + s_j - 1} \left(a_j \prod_{i=1, i\neq j}^{n} p_i^{t_i + s_i} \alpha \right),$$

so that $i(\eta_j, \Xi_1) \geqq s_j$. Thus

$$i(\eta, \Xi_1) \leqq \sum_{j=1}^{n} i(\eta_j, \Xi_1).$$

On the other hand, let $i(\eta_j, \Xi_1) = r_j$, so that

$$p_j^{t_j} \eta_j = p_j^{r_j + t_j - 1} \alpha_j, \qquad j = 1, 2, \cdots, n.$$

From (1), this implies

$$(3) \qquad a \prod_{i=1}^{n} p_i^{t_i - 1} \eta = \sum_{j=1}^{n} p_j^{r_j + t_j - 1} \alpha_j, \qquad a = \sum_{j=1}^{n} a_j \prod_{i=1, i\neq j}^{n} p_i.$$

As $(a, p_j) = 1$, $j = 1, 2, \cdots, n$, $a\bar{a} + \bar{p} p_1 p_2 \cdots p_n = 1$ for some $\bar{a}$, $\bar{p} \in Q$. If we define c_j and d_j, $j = 1, 2, \cdots, n$, as solutions of the equations

$$c_j \prod_{i=1, i\neq j}^{n} p_i^{r_i + t_i} + d_j p_j^{t_j} = 1, \qquad j = 1, 2, \cdots, n,$$

then for $\beta_j = c_j \alpha_j$, $j = 1, 2, \cdots, n$ it follows that

$$\alpha_j = \prod_{i=1, i\neq j}^{n} p_i^{r_i + t_i} \beta_j, \qquad j = 1, 2, \cdots, n.$$

A substitution of this in (3) yields

$$\prod_{i=1}^{n} p_i^{t_i - 1} \eta = \prod_{i=1}^{n} p_i^{r_i + t_i - 1} \left(\bar{a} \sum_{j=1}^{n} \prod_{i=1, i\neq j}^{n} p_i \alpha_j \right),$$

so that

$$i(\eta, \Xi_1) \geqq \sum_{j=1}^{n} i(\eta_j, \Xi_1).$$

In case $i(\eta, \Xi_1) = \infty$, it will be possible to find an α in (2) for which s exceeds any given number. Thus it must be possible to find an α for which one of the s_j exceeds any given number. This implies $i(\eta_j, \Xi_1) = \infty$ for some value j. Thus the theorem is seen to hold in all cases.

If $\eta \in H_p$, p a prime, and η minimally annihilated by p^t, then for any $a \in Q$, $a = p^r a_1$ with $(a_1, p) = 1$,

$$i(a\eta, H_p) = i(\eta, H_p) + r$$

in case $r < t$. This immediately leads to the following

Corollary 4.4. *For $\eta \in \Xi_1 \wedge H$, η minimally annihilated by $\prod_{i=1}^{n} p_i^{t_i}$, and $a = a_1 \prod_{i=1}^{m} p_i^{r_i}$, a_1 a unit, $m \geqq n$, and $a\eta \neq 0$,*

$$i(a\eta, \Xi_1) = i(\eta, \Xi_1) + \sum_{j=1, r_j < t_j}^{n} r_j.$$

For a regular element ξ of Ξ_1,

$$i(a\xi, \Xi_1) = i(\xi, \Xi_1) + \sum_{j=1}^{m} r_j.$$

The importance of the concepts of primitivity and index in connection with our problem is seen in the next theorem.

Theorem 4.5. *Let the Q-module Ξ_2 be primitive in the Q-module Ξ_1, and either $\Xi_1 \subset H_p$ for some prime p or Ξ_1 be regular. Then, for any element ξ of $\Xi_1 - \Xi_2$ such that $i(\xi, \Xi_1/\Xi_2) = 0$ and ξ has the same minimum annihilator* mod Ξ_2 *and* mod 0, *$\Xi_2 + Q\xi$ is primitive in Ξ_1.*

In the first place, if $\Xi_1 \subset H_p$, then ξ will have some power of p, say p^t, as minimum annihilator. By assumption, $a\xi \equiv 0$ mod Ξ_2 implies $a\xi = 0$, so that the sum $\Xi_2 + Q\xi$ is supplementary. Let p^s be the minimum annihilator mod $(\Xi_2 + Q\xi)$ of an element $\eta \in \Xi_1$, with $p^s\eta \not\equiv 0$ mod Ξ_2. Thus $p^s\eta \equiv a\xi$ mod Ξ_2 for some a $\in Q$, so that $i(a\xi, \Xi_1/\Xi_2) \geqq s$. If $a = bp^r$, $(b, p) = 1$, then $i(a\xi, \Xi_1/\Xi_2) = r$ by Corollary 4.4. Thus $s \leqq r$, so that $p^s(\eta - p^{r-s}b\xi) \equiv 0$ mod Ξ_2. As Ξ_2 is primitive in Ξ_1, there must exist an element ξ_1 of Ξ_1 for which $p^s\xi_1 = 0$ and $\eta - p^{r-s}b\xi + \xi_1 \in \Xi_2$. This last can be written $\eta + \xi_1 \in \Xi_2 + Q\xi$, which establishes the primitivity of $\Xi_2 + Q\xi$ in Ξ_1.

In case Ξ_1 is regular, let η be a nonzero element of Ξ_1 for which $a\eta \equiv 0$ mod $\Xi_2 + Q\xi$, $a \neq 0$. This means that there exists an element $b \in Q$ such that $a\eta \equiv b\xi$ mod Ξ_2. If $(a, b) = d = ra + sb$, $b = db_1$, $a = da_1$, then let $\xi_2 = r\xi + s\eta$. We then see that $a\xi_2 \equiv d\xi$ mod Ξ_2, or, because of the primitivity of Ξ_2 in Ξ_1,

$\xi \equiv a_1\xi_2 \bmod \Xi_2$. As $i(\xi, \Xi_1/\Xi_2)=0$, a_1 must be a unit of Q so that $a|b$ and $\eta \equiv a_1^{-1}b_1\xi \bmod \Xi_2$. Thus $\eta \in \Xi_2+Q\xi$ and the theorem is established.

Now suppose Ξ_1 is a Q-module, and Ξ_1/H_1 has a proper Q-basis, H_1 being the set of all non-regular elements of Ξ_1. Let this basis be $(\xi_1+H_1, \xi_2+H_1, \cdots)$. If $\sum_{i=1}^n a_i\xi_i \equiv 0 \bmod H_1$, then $a_i\xi_i \equiv 0 \bmod H_1$, $i=1, 2, \cdots, n$, so that $\xi_i=0$, $i=1, 2, \cdots, n$. Thus

$$\Xi_1 = H_1 + Q\xi_1 + Q\xi_2 + \cdots,$$

and the following theorem is seen to hold in view of Theorem 1.4.

THEOREM 4.6. *If Ξ_1 is a Q-module, then Ξ_1 has a proper Q-basis if and only if*

(1) *each nonzero $\Xi_1 \wedge H_{p_i}$ has a proper Q-basis, $i=1, 2, \cdots$, and*

(2) *$\Xi_1/\Xi_1 \wedge H$ has a proper Q-basis in case $\Xi_1 \neq \Xi_1 \wedge H$.*

THEOREM 4.7. *If Ξ_1 has a proper Q-basis, then all the elements of Ξ_1 have finite index in Ξ_1.*

To establish this, suppose first that $i(\eta, \Xi_1)=\infty$ for an element $\eta \in H \wedge \Xi_1$. From Theorem 4.3, we see that there must exist an integer r and an element $\alpha \in H_{p_r} \wedge \Xi_1$ such that $i(\alpha, H_{p_r} \wedge \Xi_1)=\infty$. Let $H_{p_r} \wedge \Xi_1 = Q\eta_1+Q\eta_2+\cdots$, and $\alpha=\sum_{i=1}^n a_i\eta_i$. By assumption, if α has minimum annihilator p_r^t, there exist elements $\alpha_j \in H_{p_r} \wedge \Xi_1$ and integers $t_1<t_2<\cdots$ such that

$$p_r^{t-1}\alpha = p_r^{t_j}\alpha_j, \qquad j=1, 2, \cdots.$$

If $\alpha_j=\sum_{i=1}^{n_j} a_{ji}\eta_i$, then

$$p_r^{t-1}a_i\eta_i = p_r^{t_j}a_{ji}\eta_i, \qquad i=1, 2, \cdots, n; j=1, 2, \cdots.$$

However, each η_i is annihilated by some power of p_r: therefore this last equation implies that $p_r^{t-1}\alpha=0$, which contradicts our assumptions. Thus no non-regular element of Ξ_1 can have infinite index.

Now assume $i(\xi, \Xi_1)=\infty$ for a regular element ξ of Ξ_1. Let a proper Q-basis of Ξ_1 be $(\eta_1, \eta_2, \cdots, \xi_1, \xi_2, \cdots)$, with the $\eta_i \in H$ and the ξ_i regular. Thus $p\xi = p\sum_{i=1}^n a_i\xi_i$, $p\neq 0$. By assumption, there must exist nonzero elements q_j, $r_j \in Q$ and $\beta_j \in \Xi_1$ such that $r_j\xi = r_jq_j\beta_j$, $j=1, 2, \cdots$, with the number of prime factors of q_j increasing with j. If

$$r_j\beta_j = r_j\sum_{i=1}^{n_j} a_{ji}\xi_i, \qquad j=1, 2, \cdots,$$

then

$$pr_j\sum_{i=1}^{n} a_i\xi_i = pq_jr_j\sum_{i=1}^{n_j} a_{ji}\xi_i, \qquad j=1, 2, \cdots.$$

Thus

$$pr_j(a_i - q_ja_{ji})\xi_i = 0, \qquad i = 1, 2, \cdots, n; j = 1, 2, \cdots,$$

so that

$$a_i = q_ja_{ji}, \quad i = 1, 2, \cdots, n; j = 1, 2, \cdots:$$

If a_i has t_i prime factors, then j can be taken so large that q_j has more than t_i prime factors, $i=1, 2, \cdots, n$. This is impossible, so that no regular element of Ξ_1 can have infinite index. This establishes the theorem.

THEOREM 4.8. *Let H_1 be a Q-module with the property that every element of H_1 is annihilated by some power of the prime p. Then H_1 has a proper Q-basis if and only if for every primitive set $H_2 \subset H_1$, H_2 having a finite Q-basis, all the elements of H_1/H_2 have finite index.*

To prove the necessity of this condition, let $H_1 = Q\eta_1 + Q\eta_2 + \cdots$ and define $H_1^k = Q\eta_1 + Q\eta_2 + \cdots + Q\eta_k$. For any Q-submodule H_2 of H_1 which possesses a finite Q-basis, there exists an integer n such that $H_2 \subset H_1^n$. Thus

$$H_1^n/H_2 = Q\alpha_1 + Q\alpha_2 + \cdots + Q\alpha_m$$

by Theorem 3.4. This shows that

$$H_1/H_2 = Q\alpha_1 + Q\alpha_2 + \cdots + Q\alpha_m + Q\eta_{n+1} + Q\eta_{n+2} + \cdots,$$

and thus all the elements of H_1/H_2 have finite index by Theorem 4.7.

To establish the converse, let $(\eta_1, \eta_2, \cdots)$ be any Q-basis of H_1. Assume that Γ^m is a Q-module primitive in H_1, Γ^m has a proper Q-basis $(\alpha_1, \alpha_2, \cdots, \alpha_m)$, and $i(\alpha_j, H_1/\Gamma^{j-1}) = 0$, $j = 1, 2, \cdots, m$ with $\Gamma^0 = 0$. Also assume

$$Q[\eta_1, \eta_2, \cdots, \eta_{k-1}] \subset \Gamma^m \subset H_1.$$

If $i(\eta_k, H_1/\Gamma^m) = r > 0$, with p^{t_1} the minimum annihilator of η_k mod Γ^m, then there exists an $\alpha_{m+1} \in H_1$ such that

$$p^{t_1-1}\eta_k \equiv p^{r+t_1-1}\alpha_{m+1} \bmod \Gamma^m, \qquad p^{r+t_1}\alpha_{m+1} = 0.$$

Thus $i(\alpha_{m+1}, H_1/\Gamma^m) = 0$, and by Theorem 4.5, $\Gamma^{m+1} = \Gamma^m + Q\alpha_{m+1}$ is primitive in H_1. If p^{t_2} is the minimum annihilator of η_k mod Γ^{m+1}, then $t_2 < t_1$. As above, there exists an $\alpha_{m+2} \in H_1$, $i(\alpha_{m+2}, H_1/\Gamma^{m+1}) = 0$, and $\Gamma^{m+1} + Q\alpha_{m+2}$ is primitive in H_1. There will exist an integer n and elements $\alpha_{m+3}, \alpha_{m+4}, \cdots, \alpha_{m+n} \in H_1$, with $\Gamma^{m+n} = Q\alpha_1 + Q\alpha_2 + \cdots + Q\alpha_{m+n}$, $i(\alpha_j, H_1/\Gamma^{j-1}) = 0$, $j = 1, 2, \cdots, m+n$, and Γ^{m+n} primitive in H_1, such that

$$Q[\eta_1, \eta_2, \cdots, \eta_k] \subset \Gamma^{m+n} \subset H_1.$$

Thus, by induction, there exist elements $\alpha_1, \alpha_2, \cdots \in H_1$ such that

$$H_1 = Q\alpha_1 + Q\alpha_2 + \cdots.$$

COROLLARY 4.9. *If the Q-module Ξ_1 is annihilated by a nonzero element $h \in Q$, then Ξ_1 has a proper Q-basis.*

If $h=\prod_{i=1}^{n} p_i^{t_i}$, then

$$\Xi_1 = H_1 + H_2 + \cdots + H_n, \qquad p_i^{t_i} H_i = 0, \qquad i = 1, 2, \cdots, n.$$

Thus for any Q-module $H_i' \subset H_i$, all the elements of H_i/H_i' have finite index, and the corollary follows from Theorem 4.8.

THEOREM 4.10. *If $\Xi_1 \neq \Xi_1 \wedge H$, then $\Xi_1/\Xi_1 \wedge H$ has a proper Q-basis if and only if, for every regular Q-module $K \subset \Xi_1$ such that K has a finite Q-basis and $H \wedge \Xi_1 + K$ is primitive in Ξ_1, all elements of $\Xi_1/\Xi_1 \wedge H + K$ have finite index.*

The proof of this theorem is similar to that of the last theorem. If $\Xi_1 \wedge H = H_1$, and Ξ_1 has a proper Q-basis, then $\Xi_1 = H_1 + \Omega_1$, Ω_1 a regular Q-module. Let $(\xi_1, \xi_2, \cdots)$ be a proper Q-basis for Ω_1. For any regular Q-module K which possesses a finite Q-basis there exists an integer n such that

$$H_1 + K \subset H_1 + Q\xi_1 + Q\xi_2 + \cdots + Q\xi_n.$$

As in Theorem 4.8, we see that $\Xi_1/H_1 + K$ has a proper Q-basis, and thus all its elements have finite index.

Conversely, let $\Xi_1 = Q[\beta_1, \beta_2, \cdots] \vee H_1$, all β_i being regular. Assume that we have found regular elements $\xi_1, \xi_2, \cdots, \xi_m$ in Ξ_1 which are Q-linearly independent such that, if $\Gamma^n = Q\xi_1 + Q\xi_2 + \cdots + Q\xi_n + H_1$,

$$(1) \qquad i(\xi_j, \Xi_1/\Gamma^{j-1}) = 0, \qquad j = 1, 2, \cdots, m,$$

and

$$(2) \qquad Q[\beta_1, \beta_2, \cdots, \beta_{k-1}] \vee H_1 \subset \Gamma^m \subset \Xi_1.$$

Thus Γ^m is primitive in Ξ_1. If $i(\beta_k, \Xi_1/\Gamma^m) \neq 0$, then there exists an element $\xi_{m+1} \in \Xi_1$ such that

$$\beta_k \equiv a\xi_{m+1} \bmod \Gamma^m, \qquad i(\xi_{m+1}, \Xi_1/\Gamma^m) = 0.$$

Thus

$$Q[\beta_1, \beta_2, \cdots, \beta_k] \vee H_1 \subset \Gamma^{m+1} \subset \Xi_1,$$

and the theorem follows by induction.

It is always desirable to have the important properties of any set carry over to "admissible" subsets. In this case, the property of a module having a proper basis should carry over to submodules. This was seen to be the case for P-modules in Theorem 1.1. That such is also the case for Q-modules is demonstrated in the next theorem.

THEOREM 4.11. *If Ξ_1 has a proper Q-basis, then any Q-submodule Ξ_2 of Ξ_1 also has a proper Q-basis.*

This is a consequence of Theorems 4.8 and 4.10. Let $H_j = H_{p_1} \wedge \Xi_j$, $j = 1, 2$, and assume $H_2 \neq 0$. Then for any primitive set $H_3 \subset H_2$, H_3 having a finite Q-basis, and any element $\eta \in H_2$, $i(\eta, H_2/H_3) \leqq i(\eta, H_1/H_3)$. Now, even though H_3 need not be primitive in H_1, $i(\eta, H_1/H_3) < \infty$. Thus $i(\eta, H_2/H_3) < \infty$ so that H_2 has a proper Q-basis. A similar argument also shows that $\Xi_2/H \wedge \Xi_2$ has a proper Q-basis, so that Ξ_2 has a proper Q-basis by Theorem 4.6.

5. **The modular case.** The statement that Q is an operator domain of Ξ carries with it the assumption that $a\xi = b\xi$ for all $\xi \in \Xi$ implies $a = b$. We will now consider the case in which $a\xi = b\xi$ for all $\xi \in \Xi$ implies $a \equiv b \bmod h$, h not a unit. This is equivalent to the statement that $Q/(h)$ is an operator domain of Ξ. Let $R = Q/(h)$ and $h = \prod_{i=1}^{t} p_i^{t_i}$, p_i primes of Q and $t_i \geqq 1$, $i = 1, 2, \cdots, t$. As above, $P \subset R$, and the Fundamental Axiom will still be assumed.

If h is a prime, then R is a field and Ξ is a regular R-module. Otherwise, it is apparent that the minimum annihilators of the non-regular elements of Ξ are divisors of h. In this case, if ξ is regular, $p_1\xi$ is non-regular, so that in view of Theorem 1.4 we have

$$\Xi = H_{p_1} + H_{p_2} + \cdots + H_{p_t}. \tag{1}$$

THEOREM 5.1. *The set Ξ has a proper R-basis.*

If R is a field, this follows from Theorem 1.1. Otherwise, the proof will be to show that H_{p_1} has a proper R-basis, which will lead to a proper R-basis for Ξ in view of (1). For simplification, let $H_{p_1} = H$, $t_1 = n$, $p_1 = p$, and define H_i to be the maximum submodule of H annihilated by p^i, $0 \leqq i \leqq n$, $H_0 = 0$. Thus $H_0 \subset H_1 \subset \cdots \subset H_n = H$.

The R-module H_m/H_{m-1}, $0 < m \leqq n$, is annihilated by p. Let $(\xi_1, \xi_2, \cdots)$[4] be a P-basis for this R-module. Now discard all $\xi_i \equiv 0 \bmod R[\xi_1, \xi_2, \cdots, \xi_{i-1}] \vee H_{m-1}$. If we label the remaining set $(\bar\xi_1, \bar\xi_2, \cdots)$, then this set is a proper R-basis for H_m/H_{m-1} and is R-linearly independent in H_m. For, if

$$p^{m-1}\sum_{i=1}^{k} a_i\bar\xi_i = 0, \qquad a_k \not\equiv 0 \bmod p,$$

then $a_k\bar\xi_k \equiv 0 \bmod R[\bar\xi_1, \bar\xi_2, \cdots, \bar\xi_{k-1}] \vee H_{m-1}$. However, as a_k has an inverse mod p, this implies that $\bar\xi_k \equiv 0 \bmod R[\bar\xi_1, \bar\xi_2, \cdots, \bar\xi_{k-1}] \vee H_{m-1}$, which is impossible in view of the method of selection of $\bar\xi_k$.

Assume that H/H_{m+1} has a proper R-basis $(\alpha_1, \alpha_2, \cdots)$ for some integer m, $0 \leqq m < n-1$, this basis being a R-linearly independent set in H. Let $(\beta_1, \beta_2, \cdots)$ be a proper R-basis for H_{m+1}/H_m. Now discard all β_j such that

$$\beta_j \equiv 0 \bmod R[\alpha_1, \alpha_2, \cdots] \vee H_m \vee R[\beta_1, \beta_2, \cdots, \beta_{j-1}].$$

Denote those remaining by $(\gamma_1, \gamma_2, \cdots)$. From the way that the set

---

[4] The basis is in reality $(\xi_1 + H_{m-1}, \xi_2 + H_{m-1}, \cdots)$—see the last paragraph of §1.

$(\alpha_1, \alpha_2, \cdots, \gamma_1, \gamma_2, \cdots)$ was chosen, it can be seen that it is a proper R-basis for H/H_m. To show that this set is R-linearly independent in H, suppose

$$p^{m+1}\sum_{i=1}^{r} a_i\alpha_i = p^m \sum_{i=1}^{k} c_i\gamma_i,$$

where c_k is not congruent to zero mod p. Then c_k has an inverse mod p, so that

$$\gamma_k \equiv 0 \bmod R[\alpha_1, \alpha_2, \cdots] \vee H_m \vee R[\gamma_1, \gamma_2, \cdots, \gamma_{k-1}].$$

This contradicts the method of selection of γ_k. Thus, by induction, $H = H/H_0$ has a proper R-basis.

6. **Applications to infinite matrices.** The ordered set $(1, 2, \cdots, n, \cdots)$ of type ω will be denoted by Δ. In what is to follow, x will denote a commutative indeterminate over P, and $P[x]$ will denote the polynomial domain in x over P. The P-module to be used as the Ξ above is defined as follows:

DEFINITION 6.1. *Ξ is the P-module composed of all vectors $(a_i; i \in \Delta)$ with elements in P such that only a finite number of the elements in each vector are different from zero. Addition is ordinary vector addition.*

The vector $(a_i; i \in \Delta)$ with $a_j = 1$, $a_i = 0$ for $i \neq j$ is denoted by δ_j. The set of vectors $(\delta_1, \delta_2, \cdots)$ is a proper P-basis for Ξ, and thus the Fundamental Axiom is satisfied.

The total matric algebra of order n^2 over P is denoted by $\mathfrak{M}_n$. If A is an element of $\mathfrak{M}_n$, and η is an element of the total vector space $\mathfrak{V}_n$ of order n over P, then $A\eta$ (considering η as an $n \times 1$ matrix) is again an element of $\mathfrak{V}_n$. Thus A is an operator of $\mathfrak{V}_n$, and $\mathfrak{M}_n$ is the total operator domain of $\mathfrak{V}_n$.

A total operator domain exists for Ξ, and is equivalent to the ring $\mathfrak{M}_\omega$ below.

DEFINITION 6.2. *$\mathfrak{M}_\omega$ is the set of all matrices $(a_{rs}; r, s \in \Delta)$ over P with the property that the vectors $(a_{rs}; r \in \Delta)$ are in Ξ for all $s \in \Delta$.*

For $A, B \in \mathfrak{M}_\omega$, $A = (a_{rs}; r, s \in \Delta)$, $B = (b_{rs}; r, s \in \Delta)$, the sum and product of these are defined as usual—that is,

$$AB = \left(\sum_i a_{ri}b_{is}; r, s \in \Delta\right), \qquad A + B = (a_{rs} + b_{rs}; r, s \in \Delta).$$

Under these operations of (finite) sum and product, $\mathfrak{M}_\omega$ is a ring. The element $I = (a_{rs}; r, s \in \Delta)$, with $a_{rs} = 0$ for $r \neq s$, and $a_{rr} = 1$ is the unit element of $\mathfrak{M}_\omega$. The notation above will be simplified by omitting the range of the indices when there is no chance of ambiguity. The notation $\sum_i$ means that the summation is taken over Δ.

The matrices of $\mathfrak{M}_\omega$ are left operators of Ξ under the following definition:

for $A \in \mathfrak{M}_\omega$, $\eta \in \Xi$ with $A = (a_{rs})$, $\eta = (c_i)$,

$$A\eta = \left(\sum_j a_{ij}c_j;\ i \in \Delta \right).$$

One can think of $\mathfrak{M}_\omega$ also as a vector space with elements from Ξ. If $A \in \mathfrak{M}_\omega$, $A = (a_{rs})$, then $A = (\alpha_i;\ i \in \Delta)$, where $\alpha_j = (a_{ij};\ i \in \Delta)$.

An element A of $\mathfrak{M}_\omega$ is regular in case A possesses an inverse in $\mathfrak{M}_\omega$. The ring $\mathfrak{M}_\omega$ possesses elements which are semi-regular—that is, elements which have a right inverse but not a left, or vice versa. Such an element is N defined below (Definition 6.10). The element A is algebraic in case there exists a nonzero $m(x) \in P[x]$ such that $m(A) = 0$. Thus, if A is algebraic, there exists a polynomial $h(x)$ of minimal degree, called the minimal polynomial of A, such that $h(A) = 0$.

For any non-algebraic element A of $\mathfrak{M}_\omega$, the polynomial domain $P[A]$ is a principal ideal ring. In what is to follow, these principal ideal rings correspond to the ring Q used above. In case A is algebraic with minimal polynomial $h(x)$, then $P[A]$ is isomorphic to the ring $P[x]/h(x)$, and the theory of §5 is applicable.

THEOREM 6.3. *If* $(\xi_1, \xi_2, \cdots)$ *is a proper P-basis for* Ξ, *then* $C = (\xi_i;\ i \in \Delta)$ *is a regular element of* $\mathfrak{M}_\omega$.

To prove this, we see that there must exist $a_{ij} \in P$ such that $\sum_i a_{ij}\xi_i = \delta_j$. Let $A = (a_{rs})$; then $CA = I$, so that A is a right inverse of C. Now $CAC - C = 0$, so $C(AC - I) = 0$. If $(AC - I) = (b_{rs}) \neq 0$, there exists an integer n such that $\eta = (b_{rn};\ r \in \Delta) \neq 0$. Then $C\eta = \sum_i b_{in}\xi_i = 0$, which contradicts the hypothesis that $(\xi_1, \xi_2, \cdots)$ is a proper P-basis for Ξ. Thus $AC = I$, and C is regular.

DEFINITION 6.4. *An element A of* $\mathfrak{M}_\omega$ *is said to be reducible if and only if* Ξ *has a proper* $P[A]$*-basis.*

The definition of direct sums of finite matrices (see [4, p. 237]) can be carried over to $\mathfrak{M}_\omega$. Thus an element A of $\mathfrak{M}_\omega$ is the direct sum ($\dotplus$) of the elements A_1 of $\mathfrak{M}_n$ and A_2 of $\mathfrak{M}_\omega$, $A = A_1 \dotplus A_2$, if and only if, for $A = (a_{rs})$, $A_1 = (a_{1rs})$, $A_2 = (a_{2rs})$, $a_{rs} = a_{1rs}$ for $r, s \leq n$, $a_{r+n,s+n} = a_{2rs}$ for $r, s \in \Delta$ and otherwise $a_{rs} = 0$. By iteration, the direct sum of an infinite number of finite matrices can be defined.

DEFINITION 6.5. *An element A of* $\mathfrak{M}_\omega$ *is in reduced form in case A is the direct sum of finite matrices.*

As in the finite case, two elements A and B of $\mathfrak{M}_\omega$ are similar in case there exists a regular element T of $\mathfrak{M}_\omega$ such that $B = T^{-1}AT$.

THEOREM 6.6. *If A is reducible, then any element B similar to A is also reducible.*

Let $\Xi = P[A]\xi_1 + P[A]\xi_2 + \cdots$, $B = T^{-1}AT$. Now define $\eta_n = T^{-1}\xi_n$, $n = 1, 2, \cdots$. Then $\Xi = P[B]\eta_1 + P[B]\eta_2 + \cdots$, so that B is reducible.

THEOREM 6.7. *If A is in reduced form, then A is reducible.*

That this is true is a consequence of the fact that $\mathfrak{B}_n$ has a proper $P[B]$-basis for any $B \in \mathfrak{M}_n$.

THEOREM 6.8. *If A is reducible and all the elements of Ξ are non-regular with respect to $P[A]$, then A is similar to an element B in reduced form.*

Let $\Xi = P[A]\xi_1 + P[A]\xi_2 + \cdots$, with $h_i(A)$ the minimum annihilator of ξ_i. If t_i is the degree of $h_i(x)$, then

$$h_i(x) = x^{t_i} + \sum_{j=1}^{t_i-1} a_{ij}x^j.$$

Now define

$$T = (\xi_1, A\xi_1, \cdots, A^{t_1-1}\xi_1, \xi_2, A\xi_2, \cdots, A^{t_2-1}\xi_2, \cdots),$$

so that $T \in \mathfrak{M}_\omega$. By Theorem 6.3, T is regular in $\mathfrak{M}_\omega$. If A_i is the companion matrix(5) of $h_i(x)$, so that $A_i \in \mathfrak{M}_{t_i}$, then

$$T^{-1}AT = A_1 \dotplus A_2 \dotplus \cdots,$$

and the theorem is established.

This theorem does not include the important case of algebraic matrices. However, in view of Theorem 5.1, a similar proof leads to

THEOREM 6.9. *If A is algebraic with minimal polynomial $m(x)$, then A is similar to an element B in reduced form. If $B = B_1 \dotplus B_2 \dotplus \cdots$, then each B_i is the companion matrix of some divisor of $m(x)$.*

DEFINITION 6.10. *The elements N and N^ω of $\mathfrak{M}_\omega$ are defined as follows:*

$$N = (\delta_2, \delta_3, \cdots), \qquad N^0 = 0,$$

$$N^\omega = (N\delta_1, N^2\delta_2, N^2\delta_3, N^3\delta_4, N^3\delta_5, N^3\delta_6, \cdots).$$

That powers of N are of fundamental importance in the study of reducible matrices is seen in the following

THEOREM 6.11. *If A is reducible, A not algebraic, and Ξ is a regular $P[A]$-module, then A is similar to some power of N.*

To prove this, let

$$\Xi = P[A]\xi_1 + P[A]\xi_2 + \cdots + P[A]\xi_n.$$

If $n < \infty$, let

(5) This is the $t_i \times t_i$ matrix with 1's directly below the main diagonal, $-a_{i0}, -a_{i1}, \cdots, -a_{it_i-1}$ as the last column, and 0's elsewhere.

$$T = (\xi_1, \xi_2, \cdots, \xi_n, A\xi_1, A\xi_2, \cdots, A\xi_n, A^2\xi_1, A^2\xi_2, \cdots, A^2\xi_n, \cdots).$$

Then $T^{-1}AT = N^n$. However, if $n = \infty$, let

$$T = (\xi_1, A\xi_1, \xi_2, A^2\xi_1, A\xi_2, \xi_3, \cdots).$$

Then $T^{-1}AT = N^\omega$, and the theorem follows.

The author intends to deal at greater length in a subsequent paper with matrix algebras of different order types. However, a matrix algebra of order type $\omega 2$ will be considered briefly here.

Denote the ordered set $(1, 2, \cdots, \omega+1, \omega+2, \cdots)$ of type $\omega 2$ by Δ_2. Let Ξ_2 be the P-module composed of all vectors $(a_i; i \in \Delta_2)$ over P, with only a finite number of the elements of any vector being different from zero. $\mathfrak{M}_{\omega 2}$ is the ring of all matrices $(a_{rs}; r, s \in \Delta_2)$ over P with all $(a_{rs}; r \in \Delta_2)$ in Ξ_2. Let δ_i' be the element of Ξ_2 which has 1 in the ith place and 0 elsewhere, $i \in \Delta_2$. Then $(\delta_i'; i \in \Delta_2)$ is a proper P-basis of Ξ_2.

For $A \in \mathfrak{M}_\omega$, $A' \in \mathfrak{M}_{\omega 2}$, the correspondence

$$A \leftrightarrow \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} = A'$$

defines an isomorphism between $\mathfrak{M}_\omega$ and a subring of $\mathfrak{M}_{\omega 2}$. (Under a different correspondence, $\mathfrak{M}_\omega$ and $\mathfrak{M}_{\omega 2}$ can be shown to be actually isomorphic.) Using the notation of direct sum, $A \leftrightarrow A' = A \dotplus 0$.

If A is reducible so that $\Xi = P[A]\xi_1 + P[A]\xi_2 + \cdots$, then define $\xi_j' = (\xi_j, a_{\omega+i}; i \in \Delta)$, all $a_{\omega+i} = 0$: thus $\xi_j' \in \Xi_2$. Then

$$\Xi_2 = P[A']\xi_1' + P[A']\xi_2' + \cdots + P[A']\delta_{\omega+1}' + P[A']\delta_{\omega+2}' + \cdots,$$

so that A' is also reducible.

THEOREM 6.12. *If A is reducible, then A' is similar to $B \dotplus N^k$ where $B \in \mathfrak{M}_\omega$ is in reduced form and k is an integer or ω.*

To prove this, let

$$\Xi_2 = P[A']\eta_1' + P[A']\eta_2' + \cdots + P[A']\xi_1' + P[A']\xi_2' + \cdots + P[A']\xi_k',$$

the ξ_i' being regular and the η_i' non-regular, with $h_i(x)$ the minimum annihilator of η_i'. The degree of $h_i(x)$ is labeled n_i. Then $T' \in \mathfrak{M}_{\omega 2}$ can be chosen (assuming k finite) as

$$T' = (\eta_1', A'\eta_1', \cdots, A'^{n_1-1}\eta_1', \eta_2', A'\eta_2', \cdots, A'^{n_2-1}\eta_2', \cdots, \xi_1', \xi_2', \cdots, \xi_k',$$
$$A'\xi_1', A'\xi_2', \cdots, A'\xi_k', \cdots).$$

Theorem 6.3 is seen to carry over for Ξ_2, and thus

$$T'^{-1}A'T' = B \dotplus N^k,$$

B reduced. If $k = \infty$, the ξ_i' in T' can be arranged as in Theorem 6.11.

From Theorem 2.2, we see that every reducible matrix A of $\mathfrak{M}_\omega$ has an associated characteristic matrix. The characteristic matrix of a reducible matrix is of importance in determining the similarity of matrices, as the following theorem shows.

THEOREM 6.13. *Two reducible matrices A and B of $\mathfrak{M}_\omega$ are similar if and only if their characteristic matrices are equal.*

To establish this, first assume that $\Xi = P[A]\xi_1 + P[A]\xi_2 + \cdots$ and $B = T^{-1}AT$. Then, by Theorem 6.6, $\Xi = P[B]\eta_1 + P[B]\eta_2 + \cdots$, where $\eta_i = T^{-1}\xi_i$. For any $m(x) \in P[x]$, $m(A)\xi_i = Tm(B)\eta_i$, so η_i is regular with respect to $P[B]$ if ξ_i is regular with respect to $P[A]$. Also, if $m(x)$ is the minimum annihilator of ξ_i with respect to $P[A]$, $m(x)$ is the minimum annihilator of η_i with respect to $P[B]$. Thus the characteristic matrix of A equals the characteristic matrix of B.

Now assume A and B have the same characteristic matrix. For any proper $P[A]$-basis $(\xi_1, \xi_2, \cdots, \eta_1, \eta_2, \cdots)$ of Ξ, ξ_i regular and η_i minimally annihilated by $m_i(x)$, n_i being the degree of m_i, there exists a proper $P[B]$-basis $(\bar{\xi}_1, \bar{\xi}_2, \cdots, \bar{\eta}_1, \bar{\eta}_2, \cdots)$ of Ξ with $\bar{\xi}_i$ regular and $\bar{\eta}_i$ minimally annihilated by $m_i(x)$. There is a 1-1 correspondence $\bar{\xi}_i \leftrightarrow \xi_i$, $\bar{\eta}_i \leftrightarrow \eta_i$ between these two bases. Then let

$$T = (\xi_1, \eta_1, A\eta_1, \cdots, A^{n_1-1}\eta_1, A\xi_1, \xi_2, \eta_2, A\eta_2, \cdots, A^{n_2-1}\eta_2, A^2\xi_1, A\xi_2, \xi_3, \cdots),$$

and S be the same as T with A replaced throughout by B. Then $T^{-1}AT = S^{-1}BS$, so A and B are similar, and the theorem is established.

THEOREM 6.14. *If A is reducible and regular, then A is similar to a matrix in reduced form.*

If Ξ possesses a regular vector with respect to $P[A]$, then any proper $P[A]$-basis of Ξ must have a regular element ξ. Select T as in the proof of Theorem 6.13 with ξ the first vector of T. Then $T^{-1}AT$ has its first row (or column) composed of zeros. This implies $T^{-1}AT$ is not regular which means A is not regular. Thus Ξ can have only non-regular elements with respect to $P[A]$, and the theorem follows from Theorem 6.7.

As an example of the reduction of an element of $\mathfrak{M}_\omega$ to reduced form, take

$$A = (\delta_1 + \delta_3, -\delta_3, \delta_2 - \delta_3, -\delta_3 + \delta_4, -\delta_3 + \delta_5, -\delta_3 + \delta_6, \cdots).$$

It can be verified that A is algebraic with minimal equation $x^3 - 1 = 0$. Then $\Xi = P[A]\delta_1 + P[A](\delta_1 + \delta_4) + P[A](\delta_1 + \delta_5) + \cdots$, with $A^3 - 1$ the minimum annihilator of δ_1 and $A - 1$ the minimum annihilator of $\delta_1 + \delta_i$, $i = 4, 5, \cdots$. Then T can be chosen as

$$T = (\delta_1, \delta_1 + \delta_3, \delta_1 + \delta_2, \delta_1 + \delta_4, \delta_1 + \delta_5, \cdots)$$

$$T^{-1} = (\delta_1, -\delta_1 + \delta_3, -\delta_1 + \delta_2, -\delta_1 + \delta_4, -\delta_1 + \delta_5, \cdots)$$

so that

$$T^{-1}AT = (\delta_2, \delta_3, \delta_1, \delta_4, \delta_5, \delta_6, \cdots).$$

Bibliography

1. C. Chevalley, *L'arithmetique dans les algebres de matrices*, Paris, 1936.
2. M. H. Ingraham, *A general theory of linear sets*, Trans. Amer. Math. Soc. vol. 27 (1925) pp. 163-196.
3. M. H. Ingraham and M. C. Wolf, *Relative linear sets and similarity of matrices whose elements belong to a division algebra*, ibid. vol. 42 (1937) pp. 16–31.
4. C. C. MacDuffee, *Introduction to abstract algebra*, New York, 1940.
5. H. Zassenhaus, *Lehrbuch der Gruppentheorie*, Leipzig, 1937.

University of Wisconsin,
Madison, Wis.

FOUNDATIONS OF A GENERAL THEORY OF BIRATIONAL CORRESPONDENCES

BY

OSCAR ZARISKI

In our papers dealing with the reduction of singularities of an algebraic surface (see [8, 11]), we were forced to devote a good deal of space to certain properties of birational correspondences for which we could find no general proofs in the literature. These properties were of a general character and therefore could not be regarded as part of the reduction proof proper, although they did play an auxiliary role in the proof. A similar situation arose in our reduction proof for three-dimensional varieties (not yet published), but in this case the amount of preliminary general material on birational correspondences used in the proof was even larger and was out of proportion to the length of the reduction proof proper. It thus became increasingly clear that the procedure of treating general questions of birational correspondences only as and when these questions come up in connection with various steps of the reduction process, could no longer be followed in the case of higher varieties. Instead it seemed necessary—and also worthwhile for its own sake—to develop *systematically* in a separate paper the fundamental concepts and theorems of the theory of birational correspondences, and to do this in as general a fashion as possible. This we propose to do in the present paper. We deal here with algebraic varieties, *with or without singularities*, over an arbitrary ground field (of characteristic zero or p).

It is difficult to say which of our results are entirely novel and which are not. Since many of the results hold only for normal varieties, they would appear to be novel inasmuch as our concept of a normal variety is new. On the other hand, most of our results were known for nonsingular models. It is perhaps correct to say that the novelty of the present investigation consists in showing that most of the known properties of birational correspondences between nonsingular varieties remain true more generally for normal varieties.

Of importance for the theory of algebraic functions over arbitrary ground fields of characteristic p is the fact that our construction of normal varieties which we gave in [7] in the case of algebraically closed ground fields of characteristic zero—and which carries over without essential modifications to arbitrary fields of characteristic zero (see II. 2)—can be extended to arbitrary fields of characteristic p (II. 2). This extension is made possible by a theorem of F. K. Schmidt [5] and by the normalization theorem of Emmy

Presented in part to the Society, December 31, 1941 under the title *Normal varieties and birational correspondences* (see [12]); received by the editors September 1, 1942.

Noether. The proof of this last theorem for arbitrary ground fields (and not only for fields containing "sufficiently many" elements; see Krull [2, pp. 41–42]) is given in II.2.

A feature of the treatment is our use of valuation theory. Our very definition of a birational correspondence (II.1) is valuation-theoretic in character, and our proofs are naturally conditioned by this valuation-theoretic approach. The characterization of an integrally closed ring as intersection of valuation rings, and the ideal theory in such a ring, also play an important role in our treatment.

Thanks are due to Irvin Cohen for valuable assistance lent by him during the preparation of this paper.

The following list of contents will give an idea of the individual topics treated.

CONTENTS

PART I. VALUATION-THEORETIC PRELIMINARIES

1. Homogeneous ideals. Let $\eta_0, \eta_1, \cdots, \eta_n$ be the homogeneous coordinates of the general point of an irreducible r-dimensional algebraic variety V immersed in an n-dimensional projective space. The coordinates η_i are defined to within a linear homogeneous nonsingular transformation with coefficients in the ground field K. By the very definition of homogeneous coordinates (Zariski [7, p. 284]), if $f(\eta)=0$ is an algebraic relation between the η's over K, and if we write f as a sum of forms $f_i(\eta)$ of different degrees, then each $f_i(\eta)$ individually is zero. This is equivalent to saying that the polynomials $f(y)$ in the polynomial ring $K[y_0, y_1, \cdots, y_n]$ (the y's are indeterminates) such that $f(\eta)=0$, form a *homogeneous ideal, that is, an ideal which has a base consisting of forms.*

Let P denote the ring $K[\eta_0, \eta_1, \cdots, \eta_n]$. We shall also consider homogeneous ideals in P, that is, again ideals in P which have a base consisting of forms. These are the ideals on which the homogeneous ideals of $K[y_0, \cdots, y_n]$ are mapped in the homomorphism $K[y]\sim K[\eta]$. The ring P possesses relative automorphisms τ_λ over K, where for any element $\phi(\eta)$ in P we define: $\tau_\lambda(\phi(\eta))=\phi(\lambda\eta)$, $\lambda\in K$, $\lambda\neq 0$. It is clear that if $\mathfrak{A}$ is a homogeneous ideal in P, then $\tau_\lambda(\mathfrak{A})=\mathfrak{A}$. Conversely, we have the following theorem:

THEOREM 1. *If* K *has infinitely many elements and if an ideal* $\mathfrak{A}$ *in* P *is such that* $\tau_\lambda(\mathfrak{A})\subseteq\mathfrak{A}$, *for all* λ *in* K $(\lambda\neq 0)$, *then* $\mathfrak{A}$ *is homogeneous.*

Proof. Let $\omega=f(\eta)$ be an arbitrary element of $\mathfrak{A}$ and let $f(\eta)=f_s(\eta)+f_{s+1}(\eta)+\cdots+f_m(\eta)$, where f_i is a form of degree i. We take arbitrarily in K a set of $m-s+1$ *distinct* elements $\lambda_1, \lambda_2, \cdots, \lambda_{m-s+1}$, *all different from zero.* We have, by hypothesis:

$$\tau_{\lambda_i}(\omega) = \lambda_i^s f_s(\eta) + \lambda_i^{s+1} f_{s+1}(\eta) + \cdots + \lambda_i^m f_m(\eta) \equiv 0(\mathfrak{A}).$$

These $m-s+1$ congruences imply the congruences: $f_j(\eta)\equiv 0$, $j=s, s+1, \cdots, m$. Hence $\mathfrak{A}$ is a homogeneous ideal, as was asserted.

THEOREM 2. *If* $\mathfrak{A}=[\mathfrak{q}_1, \mathfrak{q}_2, \cdots, \mathfrak{q}_h]$ *is a normal decomposition of a homogeneous ideal* $\mathfrak{A}$ *into maximal primary components, then each* $\mathfrak{q}_i$ *is either itself homogeneous or—in the case of an embedded component—can be so selected as to be homogeneous.*

Proof. We first consider the case of an infinite ground field K. Let $\mathfrak{p}_i$ be the prime ideal associated with the primary ideal $\mathfrak{q}_i$. The infinitely many automorphisms τ_λ leave $\mathfrak{A}$ invariant and hence must permute the prime ideals $\mathfrak{p}_1, \mathfrak{p}_2, \cdots, \mathfrak{p}_h$. Consequently each of these prime ideals is left invariant by *infinitely many* automorphisms τ_λ. It follows then from the proof of the preceding theorem, that the prime ideals $\mathfrak{p}_i$ of $\mathfrak{A}$ must be all homogeneous.

If $\mathfrak{p}_i$ is an isolated prime ideal of $\mathfrak{A}$, then $\mathfrak{q}_i$ is uniquely determined. Since $\tau_\lambda(\mathfrak{p}_i)=\mathfrak{p}_i$, it follows that also $\tau_\lambda(\mathfrak{q}_i)=\mathfrak{q}_i$, whence $\mathfrak{q}_i$ is homogeneous.

Let now $\mathfrak{p}_i$ be an embedded prime of $\mathfrak{A}$. We apply to $\mathfrak{q}_i$ all the automorphisms τ_λ $(\lambda\in K, \lambda\neq 0)$ and we denote by $\mathfrak{q}_i^*$ the intersection of all the ideals $\tau_\lambda(\mathfrak{q}_i)$. From the very definition of a primary ideal, it follows immediately that $\mathfrak{q}_i^*$ is a primary ideal and that $\mathfrak{p}_i$ is its associated prime ideal. Moreover, by Theorem 1, $\mathfrak{q}_i^*$ is homogeneous, and since $\mathfrak{A}\subseteq\mathfrak{q}_i^*\subseteq\mathfrak{q}_i$, we find: $\mathfrak{A}=[\mathfrak{q}_1, \mathfrak{q}_2, \cdots, \mathfrak{q}_i^*, \cdots, \mathfrak{q}_h]$, as was asserted.

Let now K be a finite field. We adjoin to K a transcendental u and we put $K^*=K(u)$, $P^*=K^*[\eta_0, \eta_1, \cdots, \eta_n]$ (we assume that u is also a transcendental with respect to the ring P). By the preceding case, our theorem holds for the ring P^* over the new ground field K^*. We can draw from this the conclusion that the theorem also holds for the ring P, provided we first es-

tablish certain relations between the ideals in P and the ideals in P^*. If $\mathfrak{B}$ is an ideal in P, it determines in P^* the *extended ideal* $P^*\cdot\mathfrak{B}$. Vice versa, every ideal $\mathfrak{C}^*$ in P^* gives rise to a *contracted ideal* in P, namely the ideal $\mathfrak{C}^*\cap P$ (but $\mathfrak{C}^*$ is not necessarily the extended ideal of its contracted ideal). Note that every element ω^* of P^* can be written in the form:

$$\omega^* = \frac{1}{g(u)}\cdot(\omega_0 u^m + \omega_1 u^{m-1} + \cdots + \omega_m), \tag{1}$$

where $g(u)\in K[u]$ and $\omega_j\in P$. The relations which we wish to establish concern the operations of extension and contraction just described, and are as follows:

a. *An ideal $\mathfrak{B}^*$ in P^* is the extension of an ideal in P if and only if the congruence $\omega^*\equiv 0(\mathfrak{B}^*)$ implies $\omega_j\equiv 0(\mathfrak{B}^*)$, for $j=0, 1, \cdots, m$, where ω^* is written in the form* (1).

b. $P^*\cdot\mathfrak{B}\cap P=\mathfrak{B}$.

c. *If* $\mathfrak{B}_1\cap\mathfrak{B}_2=\mathfrak{B}_3$, *then* $P^*\cdot\mathfrak{B}_1\cap P^*\cdot\mathfrak{B}_2=P^*\cdot\mathfrak{B}_3$. The proofs are trivial.

d. *If $\mathfrak{p}$ is a prime ideal in P, then $P^*\cdot\mathfrak{p}$ is also a prime ideal.* This assertion is essentially equivalent to the well known theorem that if R is an integral domain and if u is a transcendental with respect to R, then $R[u]$ is also an integral domain. In the present case R is the ring $P/\mathfrak{p}$.

Let now $\mathfrak{A}$ be a homogeneous ideal in P and let us consider its extended ideal $P^*\mathfrak{A}$ in P^*. It is clear that also $P^*\mathfrak{A}$ is a homogeneous ideal in P^*. Since K^* is an infinite field, we can write

$$P^*\cdot\mathfrak{A} = [\mathfrak{q}_1^*, \mathfrak{q}_2^*, \cdots, \mathfrak{q}_h^*], \tag{2}$$

where each $\mathfrak{q}_i^*$ is a homogeneous primary ideal. If we put $\mathfrak{q}_i=\mathfrak{q}_i^*\cap P$, we find, by property b:

$$\mathfrak{A} = [\mathfrak{q}_1, \mathfrak{q}_2, \cdots, \mathfrak{q}_h].$$

It is obvious that if $\mathfrak{B}^*$ is a homogeneous ideal in P^*, then $\mathfrak{B}^*\cap P$ is also a homogeneous ideal. Hence the ideals $\mathfrak{q}_i$ are all homogeneous. Since they are obviously primary ideals, our theorem follows in view of the unicity theorems concerning the decomposition of ideals into maximal primary components.

In addition to the relations a, b, c, d, we shall also have occasion to use the following property:

e. *If $\mathfrak{q}$ is a primary ideal in P and if $\mathfrak{p}$ is its associated prime ideal, then $P^*\cdot\mathfrak{q}$ is also primary and $P^*\cdot\mathfrak{p}$ is its associated prime ideal.*

To prove e, we denote by $\mathfrak{p}^*$ the prime ideal $P^*\cdot\mathfrak{p}$ and we consider any prime ideal $\mathfrak{p}_1^*$ of $P^*\cdot\mathfrak{q}$. We have: $\mathfrak{p}_1^*\supseteq P^*\mathfrak{q}\supseteq\mathfrak{q}$, whence $\mathfrak{p}_1^*\cap P\supseteq\mathfrak{q}$. But since the contraction of a prime ideal is also a prime ideal, it follows that $p_1^*\cap P\supseteq\mathfrak{p}$, whence

$$\mathfrak{p}_1^* \supseteq \mathfrak{p}^*. \tag{3}$$

Now let ω^* be an arbitrary element of $\mathfrak{p}_1^*$ and let us write ω^* in the form (1). Since $\mathfrak{p}_1^*$ is a prime ideal of $P^*\cdot\mathfrak{q}$, there exists in P^* an element ξ^* such that $\omega^*\cdot\xi^*\equiv 0\ (P^*\cdot\mathfrak{q})$, $\xi^*\not\equiv 0\ (P^*\cdot\mathfrak{q})$. Let us also write ξ^* in a form similar to (1):

$$\xi^* = \frac{1}{h(u)}\cdot(\xi_0 u^\mu + \xi_1 u^{\mu-1} + \cdots + \xi_\mu), \qquad h(u) \in K[u],\ \xi_j \in P.$$

Since $\xi^*\not\equiv 0(P^*\cdot\mathfrak{q})$, not all the elements $\xi_0, \xi_1, \cdots, \xi_\mu$ are in $\mathfrak{q}$. We may assume that $\xi_0\not\equiv 0(\mathfrak{q})$, since otherwise we can drop the term ξ_0u^μ (it is permissible to replace ξ^* by any element of P^* which is congruent to it modulo $P^*\cdot\mathfrak{q}$). Since $\omega^*\xi^*\equiv 0(P^*\mathfrak{q})$, we must have, by a and b: $\omega_0\xi_0\equiv 0(\mathfrak{q})$, and consequently $\omega_0\equiv 0(\mathfrak{p})$, since $\xi_0\not\equiv 0(\mathfrak{q})$. From this we conclude, in view of (3), that the element

$$\frac{1}{g(u)}\cdot\left[\omega_1 u^{m-1} + \omega_2 u^{m-2} + \cdots + \omega_m\right]$$

also belongs to $\mathfrak{p}_1^*$. Continuing in the same fashion with this new element of $\mathfrak{p}_1^*$, we conclude that $\omega_0, \omega_1, \cdots, \omega_m$ are all in $\mathfrak{p}$. Hence $\omega^*\in P^*\cdot\mathfrak{p}$. Since ω^* is an arbitrary element of $\mathfrak{p}_1^*$, it follows that $\mathfrak{p}_1^*\subseteq\mathfrak{p}^*$, and this, in view of (3), yields the relation: $\mathfrak{p}_1^*=\mathfrak{p}^*$. What we have shown is that $P^*\mathfrak{q}$ has only one prime ideal, namely $\mathfrak{p}^*$, and that is exactly what is asserted in e.

The relation b shows that there is a (1, 1) correspondence between the ideals $\mathfrak{B}$ in P and their extended ideals $P^*\mathfrak{B}$ in P^*, for $P^*\mathfrak{B}_1=P^*\mathfrak{B}_2$ implies $\mathfrak{B}_1=\mathfrak{B}_2$. By the property c, this correspondence is an *isomorphism* with respect to the operation $\cap$ of intersection. It is a straightforward matter to show that this correspondence is also an isomorphism with respect to the other elementary operations on ideals, namely the operation of forming the sum, the product and the quotient of two ideals:

c_1. $P^*\cdot(\mathfrak{A}, \mathfrak{B})=(P^*\mathfrak{A}, P^*\mathfrak{B})$.
c_2. $P^*\cdot(\mathfrak{A}\mathfrak{B})=P^*\mathfrak{A}\cdot P^*\mathfrak{B}$.
c_3. $P^*\cdot(\mathfrak{A}:\mathfrak{B})=P^*\mathfrak{A}:P^*\mathfrak{B}$.

We shall refer to the relations c, c_1, c_2 and c_3 as the *isomorphism relations*. The question whether any one of these relations holds, arises quite generally whenever we deal with ideals in two rings P, P^* such that P is a subring of P^*. For a general treatment of the relationship between the ideals in two such rings, see Grell [1].

The ideal $(\eta_0, \eta_1, \cdots, \eta_n)$ is referred to as the *irrelevant* prime ideal in P. Any primary ideal whose associated prime ideal is the irrelevant prime is also referred to as irrelevant. Any prime homogeneous ideal, other than the irrelevant prime, is of dimension $s+1$, $s\geqq 0$, and defines an irreducible s-dimensional subvariety of V. Two homogeneous ideals which differ only by the irrelevant component define one and the same subvariety of V.

2. **Homogeneous and nonhomogeneous coordinates.** A preference for one or the other type of coordinates is in part a matter of taste. However, it can

be claimed that in the study of properties of algebraic varieties, the use of homogeneous coordinates is indicated whenever one deals with properties *in the large*. For instance, the concept of a normal variety (Zariski [7, p. 285]) is defined essentially in terms of homogeneous coordinates. On the contrary, in questions pertaining to *local* properties it is preferable to use nonhomogeneous coordinates. Thus, if our attention is focused on a given subvariety W of V and if, say, $\eta_0 \neq 0$ on W, that is, if η_0 does not belong to the homogeneous prime ideal by which W is defined (whence W does not lie in the hyperplane $y_0=0$), then we may find it convenient to pass to the nonhomogeneous coordinates $\xi_i=\eta_i/\eta_0$, $i=1, 2, \cdots, n$. With respect to these coordinates "*W is at finite distance*"—an expression that we shall use consistently. More generally, if $l=c_0\eta_0+c_1\eta_1+\cdots+c_n\eta_n\neq 0$ on W, $c_i \in K$, and if, say $c_\alpha \neq 0$, then the quotients η_i/l, $i\neq\alpha$, may be equally well used as nonhomogeneous coordinates ξ_i of the general point of V.

It should be understood that the field Σ of rational functions on V is the field $K(\xi_1, \xi_2, \cdots, \xi_n)$ generated by the nonhomogeneous coordinates. The field $K(\eta_0, \eta_1, \cdots, \eta_n)$ is a simple transcendental extension of Σ. The field Σ consists of all quotients $f(\eta)/g(\eta)$, where f and g are forms of like degree; in other words, Σ consists of all elements of the field $K(\eta_0, \eta_1, \cdots, \eta_n)$ which are homogeneous of degree zero (Zariski [7, p. 284]).

We have also *two coordinate rings*: the ring $K[\eta_0, \eta_1, \cdots, \eta_n]$ of the homogeneous coordinates, which we have denoted by P, and—for a given choice of the nonhomogeneous coordinates ξ_i—the ring $K[\xi_1, \xi_2, \cdots, \xi_n]$ which we shall denote by $\mathfrak{o}$. In order to elicit the relationship between the ideals in $\mathfrak{o}$ and the homogeneous ideals in P, we assume for simplicity that $\xi_i=\eta_i/\eta_0$ and we consider the ring $\mathfrak{o}^*=K(\eta_0)[\xi_1, \xi_2, \cdots, \xi_n]$. Since η_0 is a transcendental with respect to Σ, the two rings $\mathfrak{o}$ and $\mathfrak{o}^*$ are in the same relationship to each other as the two rings P, P^* of the preceding section. Therefore the correspondence between the ideals in $\mathfrak{o}$ and their extended ideals in $\mathfrak{o}^*$ satisfies all the relations a–e (in which P and P^* are naturally to be replaced by $\mathfrak{o}$ and $\mathfrak{o}^*$). We shall denote by M^* the class of $\mathfrak{o}^*$-ideals which are extensions of $\mathfrak{o}$-ideals.

From the pair of rings, $\mathfrak{o}$, $\mathfrak{o}^*$ we pass to the pair of rings $\mathfrak{o}^*$, P. We have: $\mathfrak{o}^*=K(\eta_0)[\eta_1, \eta_2, \cdots, \eta_n]$. The polynomial ring $K[\eta_0]$ is a subring of P, and thus $\mathfrak{o}^*$ is a *quotient ring* of P, since $K[\eta_0]$ is at any rate closed under multiplication(1). There is therefore a (1, 1) correspondence between the ideals $\mathfrak{A}^*$ in $\mathfrak{o}^*$ and those ideals $\mathfrak{A}$ in P which are relatively prime to all elements of $K[\eta_0]$, that is, which are such that $\mathfrak{A}:\alpha=\mathfrak{A}$, for all α in $K[\eta_0]$. The correspondence is again that of extension and contraction: $\mathfrak{A}^*=\mathfrak{o}^*\mathfrak{A}$, $\mathfrak{A}=\mathfrak{A}^*\cap P$. Prime or primary ideals $\mathfrak{A}$ go, respectively, into prime or primary ideals $\mathfrak{A}^*$.

(1) For properties of quotient rings used in the text, see Grell [1] and Krull [2, p. 18]. We recall that a quotient ring is defined as follows: if R is an integral domain and if S is a subset of R which is closed under multiplication, then the quotients α/β, $\alpha, \beta \in R$, $\beta \in S$, form a ring. This is the quotient ring R_S. A special important case is the one in which S is the set-theoretic complement of a prime ideal $\mathfrak{p}$ in R, that is, $S=R-\mathfrak{p}$. In this case one writes $R_\mathfrak{p}$ instead of R_S.

The isomorphism relations c, c_1, c_2, c_3 of the preceding section, with P^* replaced by $\mathfrak{o}^*$, continue to hold(²). These are properties which hold quite generally for quotient rings.

It is immediately seen that an $\mathfrak{o}^*$-ideal $\mathfrak{A}^*$ is in the class M^* if and only if the corresponding ideal $\mathfrak{A}$ in P is homogeneous. It is clear that a homogeneous ideal is always relatively prime to any polynomial $f(\eta_0)$ if that polynomial contains a constant term which is different from zero. Hence, for a homogeneous ideal $\mathfrak{A}$ to be relatively prime to each element of $K[\eta_0]$, it is necessary and sufficient that it be relatively prime to η_0. An equivalent condition is that *no prime ideal of* $\mathfrak{A}$ *be a divisor of* (η_0). *We have therefore a* (1, 1) *correspondence between the homogeneous ideals in P which satisfy this last mentioned condition and the ideals in* $\mathfrak{o}$. This conclusion corresponds to the obvious geometrical fact that by passing to nonhomogeneous coordinates η_i/η_0 we lose track of all the subvarieties of V which "are at infinity," that is, which lie in the hyperplane $y_0=0$.

Concretely, the relationship between two corresponding ideals $\mathfrak{a}$ and $\mathfrak{A}$ in $\mathfrak{o}$ and P, respectively, is as follows: a form $f(\eta_0, \eta_1, \cdots, \eta_n)$ belongs to $\mathfrak{A}$ if and only if $f(1, \xi_1, \cdots, \xi_n)$ belongs to $\mathfrak{a}$; a polynomial $\phi(\xi_1, \xi_2, \cdots, \xi_n)$ belongs to $\mathfrak{a}$, if and only if there exists in $\mathfrak{A}$ a form $f(\eta_0, \eta_1, \cdots, \eta_n)$ such that $f(1, \xi_1, \cdots, \xi_n)=\phi(\xi_1, \xi_2, \cdots, \xi_n)$.

3. **The center of a valuation.** Let W be an irreducible subvariety of V. If $\xi_1, \xi_2, \cdots, \xi_n$ are nonhomogeneous coordinates with respect to which W is at finite distance, then W is given by a prime ideal $\mathfrak{p}$ in the ring $\mathfrak{o}$. By the *quotient ring* $Q(W)$ of W we mean the quotient ring $\mathfrak{o}_{\mathfrak{p}}$. If $\mathfrak{P}$ is the homogeneous prime ideal which corresponds to W in the ring P, then it is easily seen that $Q(W)$ consists of all quotients $f(\eta)/g(\eta)$, where f and g are forms of like degree in $\eta_0, \eta_1, \cdots, \eta_m$ and where $g(\eta)\not\equiv 0(\mathfrak{P})$. This shows, incidentally, that $Q(W)$ is independent of the choice of the nonhomogeneous coordinates.

The relationship between the ideals in $\mathfrak{o}$ and the ideals in $Q(W)$ is the one described in the preceding section for general quotient rings. The (1, 1) correspondence is now between the ideals in $Q(W)$ and those ideals in $\mathfrak{o}$ which are relatively prime to each element(³) of $\mathfrak{o}-\mathfrak{p}$. An ideal in $\mathfrak{o}$ satisfies this condition *if and only if each prime ideal of this ideal is a multiple of* $\mathfrak{p}$. This shows that by passing from the ring $\mathfrak{o}$ (or from the ring P) to the quotient ring $Q(W)$ we lose all irreducible subvarieties of V which do not contain W. For this reason we may regard the ideal theory of $Q(W)$ as that pertaining to the "neighborhood" of W. An important property of the quotient ring $\mathfrak{o}_{\mathfrak{p}}$ is the

---

(²) However, it is to be pointed out that the set of ideals $\mathfrak{A}$ which correspond to ideals $\mathfrak{A}^*$ in the above correspondence is not in general closed under the operations of multiplication and addition of ideals. It *is* closed under the operations of intersection and quotient formation.

(³) Quite generally, given a quotient ring R_S (see footnote 1), there is a (1, 1) correspondence between the ideals in R_S and the ideals in R which are relatively prime to each element of S.

following: *the non-units in* $\mathfrak{o}_\mathfrak{p}$ *form an ideal*([4]). This is the prime ideal in $\mathfrak{o}_\mathfrak{p}$ which corresponds to (that is, is the extension of) the prime ideal $\mathfrak{p}$ in $\mathfrak{o}$.

Let v be a non-trivial valuation of the field Σ over K ($v(c)=0$ if $c\in K$ and $c\neq 0$; $v(\omega)\neq 0$ for some ω in Σ). We consider linear forms l in $\eta_0, \eta_1, \cdots, \eta_n$, with coefficients in K, such that $v(\eta_i/l)\geqq 0$, $i=0, 1, \cdots, n$. If l_0 is one such form, we consider the homogeneous ideal $\mathfrak{P}$ generated by the forms $f(\eta)$ having the following property: if $f(\eta)$ is of degree m, then $v(f/l_0^m)>0$. The ideal $\mathfrak{P}$ is independent of l_0, since if l_1 is another linear form l, then $v(l_1/l_0)=0$. Moreover, $\mathfrak{P}$ is obviously a prime ideal, different from the irrelevant prime. It is also different from the zero ideal, since v is a non-trivial valuation. Consequently $\mathfrak{P}$ defines an irreducible proper subvariety W of V, of dimension at least 0. This subvariety W we call the *center of the valuation* v (on V).

It is clear that any linear form l for which $v(\eta_i/l)\geqq 0$, $i=0, 1, \cdots, n$, is such that $l\neq 0$ on W. Conversely, if l is a linear form in the η's and if $l\neq 0$ on W, then we must have: $v(l/l_0)=0$, whence $v(\eta_i/l)=v(\eta_i/l_0)\geqq 0$. Thus the linear forms l which played an auxiliary role in our definition of the center W of v, turn out simply to be those forms which do not vanish on that center.

In terms of nonhomogeneous coordinates the center W is obtained as follows. Let us consider the nonhomogeneous coordinates $\xi_i=\eta_i/\eta_0$. Should the center W be at finite distance with respect to these coordinates, we must have $\eta_0\neq 0$ on W. But then, by the remark just made, $v(\eta_i/\eta_0)\geqq 0$, $i=0, 1, \cdots, n$, and *the entire coordinate ring* $\mathfrak{o}$ *must be contained in the valuation ring* R_v *of* v. Conversely, if $\mathfrak{o}\subseteq R_v$, then reversing the above reasoning we see immediately that the center W of v is at finite distance with respect to the nonhomogeneous coordinates ξ_i. If $f(\eta_0, \eta_1, \cdots, \eta_n)$ is a form of degree ν and if $f=0$ on W, then $v(f/\eta_0^\nu)>0$, whence $v(f(1, \xi_1, \cdots, \xi_n))>0$. Conversely, it is seen immediately that every polynomial in $\xi_1, \xi_2, \cdots, \xi_n$ which has positive value in v arises from a form in the η's which vanishes on W. Hence *the elements of* $\mathfrak{o}$ *which have positive value in the given valuation* v *form a prime ideal* $\mathfrak{p}$ *in* $\mathfrak{o}$, *and the center of* v *is the irreducible subvariety of* V *defined by the ideal* $\mathfrak{p}$. This conclusion holds for any choice of the nonhomogeneous coordinates, provided the corresponding coordinate ring is contained in the valuation ring R_v.

The following characterizations of the center of a valuation are useful in applications:

THEOREM 3. *An irreducible subvariety* W *of* V *is the center of a valuation* v *if and only if either one of the following conditions is satisfied*: (1) $Q(W)\subseteq R_v$ *and the non-units of* $Q(W)$ *are non-units of* R_v; (2) $Q(W)\subseteq R_v$ *and* W *is the maximal subvariety of* V *whose quotient ring is contained in* R_v.

Proof. Suppose that W is the center of v. If $f(\eta)/g(\eta)\in Q(W)$, where f and g

(4) Chain theorem rings with the property that their non-units form an ideal have been called by Krull "*Stellenringen*" (see Krull [3]). We propose the translation: "*local rings.*" The quotient ring of any irreducible subvariety of V is a local ring.

are forms of degree m, then $g(\eta)\neq 0$ on W. Therefore $v(g/l_0^m)=0$, while $v(f/l_0^m)\geqq 0$, and consequently $f(\eta)/g(\eta)\in R_v$. Moreover, if f/g is a non-unit in $Q(W)$, then $f=0$ on W. Hence $v(f/l_0^m)>0$, and consequently $v(f/g)>0$, that is, f/g is also a non-unit in R_v. This proves that condition (1) is necessary. To show the necessity and sufficiency of condition (2), let W_1 be another irreducible subvariety of V with the property: $Q(W_1)\subseteq R_v$. Let l_0 be a linear form in $\eta_0, \eta_1, \cdots, \eta_m$ which does not vanish on W_1. Then $\eta_i/l_0\in Q(W_1)\subseteq R_v$, whence $v(\eta_i/l_0)\geqq 0$, $i=0, 1, \cdots, n$. This shows that $l_0\neq 0$ on W. Let now $f(\eta)$ be any form in the η's which vanishes on W. If f is of degree m, then $v(f/l_0^m)>0$, and consequently $l_0^m/f\notin R_v$. Since, by hypothesis, $Q(W_1)\subseteq R_v$, we have *a fortiori*, $l_0^m/f\notin Q(W_1)$, whence $f=0$ on W_1. Thus we find that "$f=0$ on W" implies "$f=0$ on W_1." Therefore $W_1\subseteq W$, and this proves our assertion.

Now it follows immediately that condition (1) is also sufficient. For, if W_1 is any *proper* subvariety of the center W of v, then[5] $Q(W_1)\subset Q(W)$ and there exist non-units in $Q(W_1)$ which are units in $Q(W)$ and which are therefore also units in R_v. This completes the proof of the theorem.

We shall conclude this section with two lemmas which we shall have occasion to use in the sequel.

LEMMA 1. *If W and W_1 are irreducible subvarieties of V, then $W_1\subseteq W$ if and only if $Q(W_1)\subseteq Q(W)$, and*[6] *$W_1\subset W$ if and only if $Q(W_1)\subset Q(W)$.*

The proof is straightforward. If $W_1\subseteq W$ and if W_1 is at finite distance with respect to the nonhomogeneous coordinates ξ_i, then also W is at finite distance, and the corresponding prime ideals $\mathfrak{p}_1$ and $\mathfrak{p}$ are such that $\mathfrak{p}_1\supseteq\mathfrak{p}$. Hence $\mathfrak{o}_{\mathfrak{p}_1}\subseteq\mathfrak{o}_{\mathfrak{p}}$. If $\mathfrak{p}_1\supset\mathfrak{p}$, and if α is an element of $\mathfrak{p}_1$, not in $\mathfrak{p}$, then $1/\alpha\in\mathfrak{o}_{\mathfrak{p}}$, but $1/\alpha\notin\mathfrak{o}_{\mathfrak{p}_1}$, whence $\mathfrak{o}_{\mathfrak{p}_1}\subset\mathfrak{o}_{\mathfrak{p}}$. Conversely, assume that $Q(W_1)\subseteq Q(W)$. If $\mathfrak{P}$ and $\mathfrak{P}_1$ are the homogeneous ideals corresponding respectively to W and W_1, let $g(\eta)$ be an arbitrary form such that $g\neq 0$ on W_1. Let $f(\eta)$ be a form of the same degree as g, such that $f\neq 0$ on W. Then $f/g\in Q(W_1)$, and hence $f/g\in Q(W)$. This implies $g\neq 0$ on W, in view of the assumption that $f\neq 0$ on W. Hence if $g\neq 0$ on W_1 we must also have: $g\neq 0$ on W. This shows that $\mathfrak{P}_1\supseteq\mathfrak{P}$, whence $W_1\subseteq W$, as asserted.

LEMMA 2. *If v and v_1 are two valuations of Σ/K, with centers W and W_1, respectively, and if v is composite with v_1, then $W\subseteq W_1$.*

Proof. A valuation v is composite with another valuation v_1, if v is obtained by combining v_1 with a valuation v' of the residue field Σ_1 of v_1. The manner in which v and v' are to be combined is best described in terms of the homomorphic mapping of Σ upon the residue field of the valuation (together with the symbol ∞), a mapping which is determined by the valuation

---

(5) See the lemma which follows immediately.

(6) We use the symbol $\subset$ only for *proper* subsets.

and which in its turn determines the valuation uniquely. Let τ_1 be the homomorphic mapping of Σ upon (Σ_1, ∞) determined by v_1, and let τ' be the homomorphic mapping of Σ_1 upon the residue field Σ_1' (and the symbol ∞) of v'. Then v is *the* valuation of Σ determined by the homomorphic mapping $\tau=\tau_1\tau'$ of Σ onto (Σ_1', ∞). For further details, see Krull [2, p. 112].

Now let $f(\eta)$ be a form which vanishes on W_1 and let $g(\eta)$ be a form of the same degree as $f(\eta)$ such that $g\neq 0$ on W *and* on W_1. The quotient $\zeta=f/g$ is a non-unit of $Q(W_1)$. Hence, by Theorem 3, ζ is also a non-unit in R_{v_1}. Consequently $\tau_1(\zeta)=0$, and therefore also $\tau(\zeta)=0$. Hence ζ is a non-unit in R_v, and, in view of our assumption that g is *not* zero on W, this is only possible if $f=0$ on W, q.e.d.

4. **Existence theorems for valuations with a preassigned center.** If R_v is the valuation ring of a given valuation v of Σ/K and if $\mathfrak{P}_v$ denotes the prime ideal of non-units of R_v, then by the dimension of v is meant the degree of transcendency of the residue field $R_v/\mathfrak{P}_v$ (over K). Let W be the center of v on V and let $\mathfrak{J}$ denote the quotient ring $Q(W)$. By Theorem 3 we have: $\mathfrak{P}_v\cap\mathfrak{J}=\mathfrak{m}$, where $\mathfrak{m}$ is the ideal of non-units in $\mathfrak{J}$. Hence $\mathfrak{J}/\mathfrak{m}$ is a subring of $R_v/\mathfrak{P}_v$, and therefore *the dimension of W is not greater than the dimension of v.* If v is of dimension $r-1$, it is called a *divisor*. A divisor is of the *first or of the second kind with respect to V*, according as its center on V is of dimension $r-1$ or less than $r-1$.

THEOREM 4. *Given an s-dimensional irreducible subvariety W of V, there exist valuations of center W, of any dimension ρ, $s\leqq\rho\leqq r-1$.*

Proof. We consider first two special cases: (a) $s=r-1$; (b) $s<r-1$, $\rho=r-1$.

Case (a) $(s=r-1)$. Let $\mathfrak{J}^*$ be the integral closure of $\mathfrak{J}$ in Σ. The $(r-1)$-dimensional prime ideal $\mathfrak{m}$ of $\mathfrak{J}$ may split in $\mathfrak{J}^*$ into several prime ideals $\mathfrak{m}_1^*, \mathfrak{m}_2^*, \cdots, \mathfrak{m}_h^*$, all of the same dimension $r-1$. It is well known that the quotient rings $\mathfrak{J}^*_{\mathfrak{m}_i^*}$ are valuation rings of divisors $v_1, v_2, \cdots, v_h$. The center of each divisor v_i is our preassigned W, and in this fashion all the divisors of center W are obtained.

Case (b) $(s<r-1, \rho=r-1)$. Assuming that W is at finite distance with respect to the nonhomogeneous coordinates ξ_i, let $\mathfrak{o}$ denote, as usual, the ring of these coordinates, and let $\mathfrak{p}$ be the prime ideal of W in $\mathfrak{o}$. Let $\omega_1, \omega_2, \cdots, \omega_m$ be a base of the ideal $\mathfrak{p}$. We select one element among these m elements ω_i and we denote it by ω. We pass to the ring $\mathfrak{o}'=\mathfrak{o}[\omega_1/\omega, \omega_2/\omega, \cdots, \omega_m/\omega]$, and we first prove the following lemma:

LEMMA 3. *For at least one mode of selecting the element ω among the elements $\omega_1, \omega_2, \cdots, \omega_m$, the following relation will be satisfied*: $\mathfrak{o}'\cdot\mathfrak{p}\cap\mathfrak{o}=\mathfrak{p}$.

Proof of the lemma. Since $\omega_i=(\omega_i/\omega)\cdot\omega\in\mathfrak{o}'\cdot\omega$, it follows that the ideal $\mathfrak{o}'\cdot\mathfrak{p}$ coincides with the principal ideal $\mathfrak{o}'\cdot\omega$. Now let us suppose that

$\mathfrak{o}' \cdot \omega \cap \mathfrak{o} \neq \mathfrak{p}$, and let us see what restriction, if any, this assumption imposes on the element ω. Let ζ be an element of $\mathfrak{o}$, which belongs to $\mathfrak{o}' \cdot \omega$ but not to $\mathfrak{p}$. We will have then: $\zeta = \omega \cdot f(\omega_1/\omega, \omega_2/\omega, \cdots, \omega_m/\omega)$, where $f(z) = f(z_1, z_2, \cdots, z_m) \in \mathfrak{o}[z_1, z_2, \cdots, z_m]$. If ν is the degree of f, the above expression for ζ leads to a relation of the form: $\zeta \cdot \omega^{\nu-1} = \phi(\omega_1, \omega_2, \cdots, \omega_m)$, where ϕ is a form of degree ν, with coefficients in $\mathfrak{o}$. This relation implies that $\zeta \cdot \omega^{\nu-1}$ is in $\mathfrak{p}^\nu$. Since $\zeta \not\equiv 0(\mathfrak{p})$, we conclude that $\mathfrak{p}^\nu : \omega_i^{\nu-1}$ *is a proper divisor of* $\mathfrak{p}$.

If our lemma is false, then for each element ω_i, $i = 1, 2, \cdots, m$, there must exist an integer ν_i such that $\mathfrak{p}^{\nu_i} : \omega_i^{\nu_i - 1} \supset \mathfrak{p}$. If $\sigma = \text{maximum}\ (\nu_1, \nu_2, \cdots, \nu_m)$, then we will have $\mathfrak{p}^\sigma : \omega_i^{\sigma-1} \supset \mathfrak{p}$, $i = 1, 2, \cdots, m$. Hence we have also $\mathfrak{p}^{\sigma+\rho} : \mathfrak{p}^\rho \omega_i^{\sigma-1} \supset \mathfrak{p}$, for any integer ρ. Now if $q = (\sigma - 2) \cdot m + 1$, then it is clear that $(\mathfrak{p}^\rho \omega_1^{\sigma-1}, \mathfrak{p}^\rho \omega_2^{\sigma-1}, \cdots, \mathfrak{p}^\rho \omega_m^{\sigma-1}) = \mathfrak{p}^q$, if $\rho = q - \sigma + 1$. Therefore the quotient $\mathfrak{p}^{q+1} : \mathfrak{p}^q$ is the intersection of the ideals $\mathfrak{p}^{q+1} : \mathfrak{p}^{q-\sigma+1} \omega_i^{\sigma-1}$, $i = 1, 2, \cdots, m$. But each of these m ideals is, by hypothesis, a proper divisor of $\mathfrak{p}$. Hence also $\mathfrak{p}^{q+1} : \mathfrak{p}^q$ is a proper divisor, and this is impossible since(7) $\mathfrak{p}^{q+1} : \mathfrak{p}^q = \mathfrak{p}$. Our lemma is thus proved.

We therefore may assume that $\mathfrak{o}' \cdot \mathfrak{p} \cap \mathfrak{o} = \mathfrak{p}$. This implies at any rate that the ideal $\mathfrak{o}' \cdot \omega$ is not the unit ideal, whence its minimal prime ideals are all $(r-1)$-dimensional. The relation $\mathfrak{o}' \cdot \mathfrak{p} \cap \mathfrak{o} = \mathfrak{p}$ also implies, and is in fact equivalent to, the assertion that at least one minimal prime ideal of $\mathfrak{o}' \cdot \omega$ must contract to $\mathfrak{p}$. Let $\mathfrak{p}'$ be such a minimal prime.

Now let V' be the projective model whose general point (in nonhomogeneous coordinates) is(8) $(\xi_1, \xi_2, \cdots, \xi_n, \omega_1/\omega, \omega_2/\omega, \cdots, \omega_m/\omega)$, so that $\mathfrak{o}'$ is the ring of the nonhomogeneous coordinates of the general point of V'. Let W' be the $(r-1)$-dimensional subvariety of V' which is defined by the prime ideal $\mathfrak{p}'$. By the preceding case (a) there exists a $(r-1)$-dimensional valuation whose center on V' is W'. Since $\mathfrak{p}' \cap \mathfrak{o} = \mathfrak{p}$, it follows that the center of v on V is W. This establishes our theorem in the case under consideration.

To prove the theorem in the general case, $s < r-1$, $s \leqq \rho < r-1$, we shall keep s and ρ fixed and we shall proceed by induction with respect to r, since, by the case (b), the theorem is true if $r = \rho + 1$. We consider an $(r-1)$-dimensional irreducible subvariety W_1 of V which contains W and we denote by v_1 a divisor of center W_1. The residue field Σ_1^* of v_1 is a finite algebraic extension of the field Σ_1 of rational functions on W_1. By our induction there exists a valuation v' of Σ_1, of dimension ρ, whose center on W_1 is W. This valuation v' has at least one extension v^* in Σ_1^*. Compounding v_1 with v^* we get a composite valuation v of Σ, of dimension ρ, whose center is W. This completes the proof of our theorem.

---

(7) Let $\mathfrak{p}^{q+1} : \mathfrak{p}^q = \mathfrak{A}$. We have then: $\mathfrak{A}\mathfrak{p}^q \equiv 0(\mathfrak{p}^{q+1})$. On the other hand $\mathfrak{p} \subseteq \mathfrak{A}$, whence $\mathfrak{A} \cdot \mathfrak{p}^q \supseteq \mathfrak{p}^{q+1}$. Consequently $\mathfrak{A} \cdot \mathfrak{p}^q = \mathfrak{p}^{q+1} = \mathfrak{p} \cdot \mathfrak{p}^q$. From this it follows (see Krull [2, p. 36]) that $\mathfrak{A}$ and $\mathfrak{p}$ have the same radical. Since $\mathfrak{A} \supseteq \mathfrak{p}$, it follows that $\mathfrak{A} = \mathfrak{p}$.

(8) Comparison with section 11 will show that our V' is the transform of V by a monoidal transformation of center W.

The preceding proof does not give an adequate idea of the totality of all valuations having a preassigned center. The valuations obtained in the course of the proof are special in the sense that if their dimension is ρ then the rank of their value group is $r-\rho$. The following theorem gives more information about the arbitrary elements which can be assigned in the construction of valuations with a given center(9):

THEOREM 5. *Given an arbitrary descending chain*(10) $W_0 \supseteq W_1 \supseteq \cdots \supseteq W_{\sigma-1}$ *of irreducible subvarieties of* V *and given any set of integers* $\rho_0, \rho_1, \cdots, \rho_{\sigma-1}$ *such that* $r-1 \geqq \rho_0 > \rho_1 > \cdots > \rho_{\sigma-1}$, $\rho_i \geqq$ *dimension of* W_i, *there exists a sequence of valuations* $v_0, v_1, \cdots, v_{\sigma-1}$ *such that*: (1) v_i *is of dimension* ρ_i, *of rank* $i+1$, *and its center is* W_i; (2) v_i *is compounded with* v_{i-1}.

We first prove the theorem in the following two special cases: (a) $\sigma=1$, $\rho_0=s=$dimension of W_0; (b) $\sigma=1$, $\rho_0>s$.

(a) $\sigma=1$, $\rho_0=s=$dimension W_0. In this case we have to prove the existence of a rank 1 valuation, of dimension s, whose center is a given s-dimensional irreducible subvariety W_0 of V. We shall use nonhomogeneous coordinates $\xi_1, \xi_2, \cdots, \xi_n$ with respect to which W_0 is at finite distance, so that W_0 is given by a prime ideal $\mathfrak{p}$ in the ring $\mathfrak{o}$ of these coordinates. We then adjoin to the ground field s elements of $\mathfrak{o}$ which are algebraically independent on W, that is, algebraically independent mod $\mathfrak{p}$. In this fashion we achieve a reduction to the case $s=0$, so that we may assume that W_0 is a point, say A. It is also permissible to assume(11) that $\xi_{r+1}, \xi_{r+2}, \cdots, \xi_n$ are

(9) Which ordered groups can be preassigned as value groups for valuations of fields of algebraic functions, is a question which has been solved completely by S. MacLane and O. F. G. Schilling in their paper [4].

(10) Note that we do not assume that the chain is strictly descending, that is, that each W_i is a proper subvariety of W_{i-1}.

(11) In the case of infinite ground fields, or of ground fields with "sufficiently many" elements, this assumption can always be realized, in view of the usual proof of Emmy Noether's normalization theorem, by first subjecting the nonhomogeneous coordinates $\xi_1, \xi_2, \cdots, \xi_n$ to a linear transformation with "non-special" coefficients in K. In the case of finite ground fields this is no longer true. However, it will be proved in II.2 that there always exists in the ring $K[\xi_1, \xi_2, \cdots, \xi_n]$ a set of r algebraically independent elements $\zeta_1, \zeta_2, \cdots, \zeta_r$ such that the ring is integrally dependent on the ring $K[\zeta_1, \zeta_2, \cdots, \zeta_r]$. We may then simply include the elements ζ_i among the elements ξ_i, which does not change the ring of nonhomogeneous coordinates, and we may then proceed as in the text.

We could also proceed in the following fashion. We first pass to the field K' which is generated over K by the coefficients of the linear transformation on the ξ_i mentioned above, and we consider an extension field $\Sigma'=K'\Sigma$ of Σ. The field K' may be assumed to be an algebraic extension of K. We obtain a new variety V' *over* K', with the same general point $(\xi_1, \xi_2, \cdots, \xi_n)$ as V. The original subvariety W_0 splits on V' into at most a finite number of varieties, all of the same dimension as W_0. If W_0' is one of them, the proof given in the text leads to a valuation v' of Σ' of dimension s and rank 1, with center W_0'. The valuation v of Σ induced by v' will have center W_0, dimension s and rank 1.

integrally dependent on $K[\xi_1, \xi_2, \cdots, \xi_r]$. Let A' be the projection([12]) of the point A into the linear space of the r independent variables $\xi_1, \xi_2, \cdots, \xi_r$. Let v' be a zero-dimensional and rank 1 valuation of the field $K(\xi_1, \xi_2, \cdots, \xi_r)$ whose center in the above linear space is the point([13]) A'. We denote, as usual, by $R_{v'}$ the valuation ring of v', and by $A_1, A_2, \cdots, A_h$ the other points of V, at finite distance and different from A, which project into A'. We can find an element ω in $\mathfrak{o}$ such that $\omega=0$ at A, $\omega\neq 0$ at A_i, $i=1, 2, \cdots, h$. Let

$$\omega^m + a_1(\xi_1, \xi_2, \cdots, \xi_r)\omega^{m-1} + \cdots + a_m(\xi_1, \xi_2, \cdots, \xi_r) = 0$$

be the irreducible equation of integral dependence for ω over the ring $K[\xi_1, \xi_2, \cdots, \xi_r]$. Since $\omega=0$ at A, we must have $a_m=0$ at A'. Hence a_m is a non-unit in $R_{v'}$. *We assert that $R_{v'}[\omega]$ is a proper ring* (that is, is not a field). We prove this by showing that *ω is a non-unit in this ring*. For suppose that ω is a unit in $R_{v'}[\omega]$. Then we would have: $1=\omega\cdot g(\omega)$, where $g(\omega)$ is a polynomial with coefficients in $R_{v'}$. Using the above equation of integral dependence and observing that the coefficients a_i are polynomials, hence are elements of $R_{v'}$, we can reduce the degree of $g(\omega)$. We thus find a new relation of the form: $1=\omega(b_0\omega^{m-1}+b_1\omega^{m-2}+ \cdots +b_{m-1})$, where $b_i\in R_{v'}$. Comparing this relation with the above relation of integral dependence, we conclude that $b_0=-1/a_m$, a contradiction, since a_m is a non-unit in $R_{v'}$.

Since ω is a non-unit in $R_{v'}[\omega]$, there exists at least one valuation v of Σ such that $R_v\supseteq R_{v'}[\omega]$ and such that ω is a non-unit([14]) in R_v. Since v' is of rank 1, $R_{v'}$ is a maximal subring of $K(\xi_1, \xi_2, \cdots, \xi_r)$. Hence $R_v\cap K(\xi_1, \xi_2, \cdots, \xi_r)=R_{v'}$, and therefore *the valuation v is an extension of v'* and has the same rank and the same dimension as v', that is, rank 1 and dimension 0. The center of v on V must be a point at finite distance (since $\xi_1, \xi_2, \cdots, \xi_r\in R_{v'}\subseteq R_v$ and hence $\mathfrak{o}\subseteq R_v$, for R_v is integrally closed), and this point must project into the point A'. The center cannot be any of the points $A_1, A_2, \cdots, A_h$, since $\omega\neq 0$ at A_i and this implies that ω is a unit in the quotient ring $Q(A_i)$, while, as we have just seen, ω is a non-unit in R_v (compare with Theorem 3). Hence the center of v is the point A, q.e.d.

(b) Let now $\sigma=1$, $\rho>s$. We refer to the case (b) of the proof of Theorem 4, where we identify the variety W with our present variety W_0. We had there an $(r-1)$-dimensional prime ideal $\mathfrak{p}'$ in $\mathfrak{o}'$ such that $\mathfrak{p}'\cap\mathfrak{o}=\mathfrak{p}$. From the existence of an $(r-1)$-dimensional prime ideal in $\mathfrak{o}'$ which contracts to $\mathfrak{p}$ follows immediately the existence of prime ideals in $\mathfrak{o}'$ of any dimension ρ,

([12]) By that we mean that A' is the point which is defined by the contraction of the prime ideal $\mathfrak{p}$ in the polynomial ring $K[\xi_1, \xi_2, \cdots, \xi_r]$.

([13]) The existence of v' is proved in the joint paper by MacLane and Schilling [4].

([14]) This is implied by the fundamental theorem on principal orders which states that an integrally closed integral domain (*not* a field) is the intersection of the valuation rings which contain it (see Krull [2, p. 111]). It is necessary only to observe that the integral closure of a *proper* ring is also a proper ring. For if a ring R is proper, it contains a non-unit α, and it is seen immediately that $1/\alpha$ cannot be integrally dependent on R.

$r-1 \geqq \rho \geqq s$, which contract to(15) $\mathfrak{p}$. Let $\mathfrak{p}''$ be such a prime ideal and let W'' be the corresponding irreducible ρ-dimensional subvariety of V'. By the preceding case (a), there exists a valuation v, of rank 1 and of dimension ρ, whose center on V' is W''. Since $\mathfrak{p}'' \cap \mathfrak{o} = \mathfrak{p}$, the center of v on V is W, q.e.d.

To prove our theorem in the general case, we first prove this lemma:

LEMMA 4. *If W and W_1 are irreducible subvarieties of V such that $W \subset W_1 \subset V$ and if v_1 is a valuation of center W_1, then there exists a valuation v of center W, which is compounded with v_1.*

Proof. Let Σ_1 be the residue field of the valuation v_1 and let τ_1 denote the homomorphic mapping of Σ onto (Σ_1, ∞) defined by v_1. Since $Q(W) \subset Q(W_1)$ (Lemma 1, I.3) and $Q(W_1) \subseteq R_{v_1}$, it follows that if we put $\mathfrak{J} = Q(W)$, then $\tau_1 \cdot \mathfrak{J} \subseteq \Sigma_1$. The elements of $\mathfrak{J}$ which are mapped into zero under τ_1 are non-units in $Q(W_1)$. Since W is a proper subvariety of W_1, there are non-units in $Q(W)$ which are units in $Q(W_1)$. *Hence $\tau_1\mathfrak{J}$ is a proper ring.* There exists then at least one valuation v' of Σ_1 such that $R_{v'} \supseteq \tau_1 \cdot \mathfrak{J}$. Let v_2 denote the valuation obtained by compounding v with v', and let W_2 be the center of v_2. By Lemma 2, I.3, we have $W_2 \subseteq W_1$. Since $R_{v'} \supseteq \tau_1 \cdot \mathfrak{J}$, it follows that $R_{v_2} \supseteq \mathfrak{J}$, whence, by Theorem 3 (I.3), we have $W \subseteq W_2$. If $W_2 = W$, then our lemma is proved ($v = v_2$). If, however, W is a proper subvariety of W_2, then we replace v_1 and W_1 by v_2 and W_2 and we repeat the above procedure. Since v_2 is of *smaller dimension* than v_1, this process cannot continue indefinitely, q.e.d.

We now are in position to complete the proof of the theorem in the general case. Since by the special cases (a) and (b) treated above the theorem is true in the case $\sigma = 1$, we assume that the theorem is true for $\sigma = h-1$ and we proceed to prove the theorem for $\sigma = h$. We can therefore assume the existence of the valuation $v_0, v_1, \cdots, v_{\sigma-2}$ and we have only to prove the existence of $v_{\sigma-1}$. The valuation $v_{\sigma-2}$ is of dimension $\rho_{\sigma-2}$ and its center is $W_{\sigma-2}$. We have to prove the existence of a valuation $v_{\sigma-1}$, of dimension $\rho_{\sigma-1}$, which is compounded with $v_{\sigma-2}$, is of rank *one higher* than the rank of $v_{\sigma-2}$ and has center $W_{\sigma-1}$. For simplicity we shall denote $W_{\sigma-2}, W_{\sigma-1}, \rho_{\sigma-2}, \rho_{\sigma-1}$ and $v_{\sigma-2}$ by W_1, W, ρ_1, ρ and v_1, respectively, so that we have now:

$$W_1 \supseteq W, \quad \rho_1 > \rho \geqq \text{dimension } W; \quad \rho_1 \geqq \text{dimension } W_1.$$

We first provide ourselves with a projective model V' on which the center W_1' of v_1 is exactly of dimension(16) ρ_1. We then consider, for auxiliary purposes,

(15) Let $\alpha_1, \alpha_2, \cdots, \alpha_s$ be s elements of $\mathfrak{o}$ which are algebraically independent modulo $\mathfrak{p}$ and let us take as new ground field the field $K_1 = K(\alpha_1, \alpha_2, \cdots, \alpha_s)$. If $\mathfrak{o}_1 = K_1 \cdot \mathfrak{o}$, $\mathfrak{o}_1' = K_1 \cdot \mathfrak{o}'$, $\mathfrak{o}_1 \cdot \mathfrak{p} = \mathfrak{p}_1$, $\mathfrak{o}_1' \cdot \mathfrak{p}' = \mathfrak{p}_1'$, then, *over* K_1, $\mathfrak{p}_1$ is zero-dimensional, $\mathfrak{p}_1'$ is $(r-1-s)$-dimensional, and $\mathfrak{p}_1' \cap \mathfrak{o}_1 = \mathfrak{p}_1$. Let $\mathfrak{p}_1''$ be any $(\rho-s)$-dimensional prime divisor of $\mathfrak{p}_1'$, and let $\mathfrak{p}_1'' \cap \mathfrak{o}' = \mathfrak{p}''$. Then $\mathfrak{p}''$ is of dimension ρ, *over* K, and $\mathfrak{p}'' \cap \mathfrak{o} = \mathfrak{p}$.

(16) As nonhomogeneous coordinates of such a model V' we may take any finite set of generators of Σ which belong to R_{v_1} and such that ρ_1 of these generators have algebraically independent residues in the residue field of v_1.

an arbitrary valuation v which is compounded with v_1 and whose center on V is W (Lemma 4; if $W_1=W$, we put $v=v_1$). Let W' denote the center of v on V', whence $W'\subseteq W_1'$ (Lemma 2, I.3). We next select nonhomogeneous coordinates $\xi_1, \xi_2, \cdots, \xi_n$ for the general point of V and nonhomogeneous coordinates $\zeta_1, \zeta_2, \cdots, \zeta_m$ for the general point of V' in such a fashion that W and W' be at finite distance with respect to these coordinates. Finally we denote by V^* the projective model whose general point is $(\xi_1, \xi_2, \cdots, \xi_n, \zeta_1, \zeta_2, \cdots, \zeta_m)$, and we denote by $\mathfrak{o}$, $\mathfrak{o}'$ and $\mathfrak{o}^*$, respectively, the corresponding rings $K[\xi]$, $K[\zeta]$, $K[\xi, \zeta]$ of the nonhomogeneous coordinates.

Let W^* and W_1^* be the centers on V^* of the valuations v and v_1, respectively. Since W and W' are at finite distance, we have: $\mathfrak{o}\subseteq R_v$, $\mathfrak{o}'\subseteq R_v$, and, by a stronger reason, $\mathfrak{o}\subseteq R_{v_1}$, $\mathfrak{o}'\subseteq R_{v_1}$. Therefore $\mathfrak{o}^*\subseteq R_v$, $\mathfrak{o}^*\subseteq R_{v_1}$, and consequently W^* and W_1^* are at finite distance. Let $\mathfrak{p}, \mathfrak{p}_1; \mathfrak{p}', \mathfrak{p}_1'; \mathfrak{p}^*, \mathfrak{p}_1^*$ denote the prime ideal of $W, W_1; W', W_1'; W^*, W_1^*$ in $\mathfrak{o}$, $\mathfrak{o}'$ and $\mathfrak{o}^*$ respectively. It is clear that:

$$\mathfrak{p}^* \cap \mathfrak{o} = \mathfrak{p}, \qquad \mathfrak{p}_1^* \cap \mathfrak{o} = \mathfrak{p}_1.$$

Moreover $W^*\subseteq W_1^*$. The above relations show that *any valuation of center* W^* *on* V^* *has* W *as center on* V. We also point out that V^* shares with V' the property that *the center of* v_1 *on that variety is exactly of dimension* ρ_1. This follows immediately from the relation: $\mathfrak{p}_1^*\cap\mathfrak{o}'=\mathfrak{p}_1'$. The variety V' and the auxiliary valuation v have now served their purpose and will not be used any more.

(1) Suppose first that W^* is of dimension at most ρ. By the special case $\sigma=1$, we can find a ρ-dimensional, rank 1 valuation *of the field of rational functions on* W_1^*, having W^* as center. This valuation has at least one extension in the residue field of the valuation v_1. Let v_2 be such an extension. Since the residue field of v_1 is an algebraic extension of the field of rational functions on W_1^*, it follows that also v_2 is of rank 1 and dimension ρ. Compounding v_1 with v_2, we get a valuation v of Σ, of dimension ρ, of rank one higher than v_1. Its center on V^* is W^*, hence its center on V is W. This valuation v is the valuation $v_{\sigma-1}$, whose existence we have claimed in our theorem. (2) Suppose now that W^* is of dimension greater than ρ. Since $\mathfrak{p}^*\cap\mathfrak{o}=\mathfrak{p}$ and since $\mathfrak{p}$ is of dimension at most ρ, it follows that we can find in $\mathfrak{o}^*$ a prime ρ-dimensional ideal which divides $\mathfrak{p}^*$ and which likewise contracts([17]) to $\mathfrak{p}$. This prime ideal defines a ρ-dimensional irreducible subvariety W_0^* of V^* which we can use with the same effect instead of W^*, since the two essential conditions: (1) $W_0^*\subseteq W_1^*$, (2) every valuation of center W_0^* has W as center on V, are still satisfied. But now W_0^* has dimension ρ, and we have therefore the case (1) just considered. This completes the proof of Theorem 5.

([17]) See footnote 15.

PART II. GENERAL THEORY OF BIRATIONAL CORRESPONDENCES

1. Valuation-theoretic definition of a birational correspondence. Let V and V' be two birationally equivalent r-dimensional irreducible algebraic varieties. The two varieties can be regarded as projective models of one and the same field Σ of algebraic functions([18]), and if they are so regarded there arises a well defined correspondence between the irreducible subvarieties of V (of all possible dimension from 0 to $r-1$ inclusive) and the irreducible subvarieties of V'. It is the *birational correspondence* between V and V', or the *birational transformation* of V into V'. This correspondence, which we shall denote by T, is defined as follows([19]):

DEFINITION 1. *Two irreducible subvarieties W and W' of V and V' respectively (not necessarily of the same dimension) correspond to each other (in symbols: $T(W)=W'$, $T^{-1}(W')=W$), if there exists a valuation v of the field Σ such that the center of v on V is W and the center of v on V' is W'.*

Note that this definition retains its full meaning also when V and V' are *coincident varieties* (as varieties in the projective space). In this case we deal with an automorphism τ of Σ and we have a *birational transformation of W into itself*. It is only necessary to regard the two coincident varieties V and V' as *distinct* projective models of Σ, in the sense that the general point of V is $(\xi_1, \xi_2, \cdots, \xi_n)$ and the general point of V' is $(\tau\xi_1, \tau\xi_2, \cdots, \tau\xi_n)$.

From the results of I.3 and I.4 we deduce immediately the following properties of a birational correspondence:

A. *Given $W \subset V$, there exists at least one $W' \subset V'$ such that $T(W)=W'$* (Theorem 4, I.4).

B. *If $W \subseteq W_1 \subset V$ and $W_1' = T(W_1)$, there exists a W' such that $W'=T(W)$ and $W' \subseteq W_1'$. In particular, if to W_1 there corresponds a point P' on V', then P' corresponds to each point of W_1.*

A birational correspondence is, generally speaking, not a (1, 1) correspondence. There may very well exist varieties W on V such that T is not single-valued at W, that is, such that to W there correspond more than one subvariety of V'. Similarly for T^{-1} and V'. These varieties W are exceptional in the sense that they lie on algebraic subvarieties of V (see Theorem 15, II.9). The analysis of these exceptions to the (1, 1) character of a birational correspondence is the main goal of our study. At this stage, however, we wish

([18]) The fields Σ, Σ' of rational functions on V and V', respectively, are isomorphic over K. When we say that V and V' are projective models of one and the same field we imply that the fields Σ, Σ' have been identified. The identification is determined to within an automorphism of Σ. When we speak of a birational correspondence we refer to a fixed identification of the two fields.

([19]) From now on *irreducible* subvarieties of V shall be denoted by the letter W, with or without subscripts. Similarly W', W_1', W_i' and so on shall always denote irreducible subvarieties of V'.

to give a very simple but important criterion for the uniqueness of $T(W)$ when W is given:

THEOREM 6. *If $T(W)=W'$ and if $Q(W')\subseteq Q(W)$, then W' is the only subvariety of V' which corresponds to W.*

Proof. By hypothesis, there exists a valuation v_1 whose center on V is W and whose center on V' is W'. Let v be an arbitrary valuation of center W. We have $R_v\supseteq Q(W)$, whence $R_v\supseteq Q(W')$. Every non-unit of $Q(W')$ is a non-unit of R_{v_1} (Theorem 3, I.3), hence it is also a non-unit of $Q(W)$. We conclude that every non-unit of $Q(W')$ is a non-unit in R_v, and our theorem follows from Theorem 3.

2. **The birational correspondence between V and a derived normal model $\bar{V}$.** In our paper [7] we have proved that from any irreducible algebraic variety V it is possible to pass to what we have called a *derived normal model $\bar{V}$ of V*. That proof dealt only with algebraically closed ground fields of characteristic zero. To extend the proof to arbitrary ground fields additional considerations are necessary.

First of all we shall need an extension of the normalization theorem of Emmy Noether to finite ground fields. Given a finite integral domain $K[\xi_1, \xi_2, \cdots, \xi_n]$ of degree of transcendency r over an infinite ground field K, the Noether normalization theorem states that there exist r linear combinations $\xi_i' = \sum_{j=1}^{n} c_{ij}\xi_j$, $i=1, 2, \cdots, r$, with coefficients in K, which are algebraically independent over K and which are such that $\xi_1, \xi_2, \cdots, \xi_n$ are integrally dependent on $\xi_1', \xi_2', \cdots, \xi_r'$. This theorem, as it stands, is not generally true when K is a finite field. In this case we can still assert that elements such as $\xi_1', \xi_2', \cdots, \xi_r'$ can be found in the ring $K[\xi_1, \xi_2, \cdots, \xi_n]$, provided we drop the condition that these elements be linear in the ξ's. The proof of this assertion, as given below, was communicated orally to me by Irvin Cohen.

We shall first consider an homogeneous integral domain $\mathfrak{o}^*=K[\eta_0, \eta_1, \cdots, \eta_n]$, of degree of transcendency $r+1$, that is, one whose generating elements η_i are the homogeneous coordinates of the general point of an r-dimensional variety V.

(a) *If the ideal $(\eta_1, \eta_2, \cdots, \eta_n)$ in $\mathfrak{o}^*$ is irrelevant, then η_0 is integrally dependent on $\eta_1, \eta_2, \cdots, \eta_n$.* For the hypothesis implies that the point $y_0=1, y_1=y_2=\cdots=y_n=0$ is not on V, and hence there must exist a form $f(y_0, y_1, \cdots, y_n)$ such that $f(\eta_0, \eta_1, \cdots, \eta_n)=0$ and $f(1, 0, \cdots, 0)\neq 0$. If ρ is the degree of f, the term η_0^ρ must therefore occur in $f(\eta)$, and this proves our assertion.

(b) More generally, *if the ideal $(\eta_{k+1}, \eta_{k+2}, \cdots, \eta_n)$ is irrelevant, then $\eta_0, \eta_1, \cdots, \eta_k$ are integrally dependent on $\eta_{k+1}, \eta_{k+2}, \cdots, \eta_n$.* Proof by induction with respect to k (that is, with respect to the number $k+1$ of elements η_i which do not occur in the set $\eta_{k+1}, \eta_{k+2}, \cdots, \eta_n$). By (a), η_0 is integrally

dependent on $\eta_1, \eta_2, \cdots, \eta_n$. Since every element of $\mathfrak{o}^*$ is integrally dependent on $\eta_1, \eta_2, \cdots, \eta_n$, the elements $\eta_{k+1}, \eta_{k+2}, \cdots, \eta_n$ generate in the ring $K[\eta_1, \eta_2, \cdots, \eta_n]$ an ideal of the same dimension as that of the ideal generated by them in $\mathfrak{o}^*$, that is, they generate in $K[\eta_1, \eta_2, \cdots, \eta_n]$ an irrelevant ideal. By our induction it follows that $\eta_1, \eta_2, \cdots, \eta_k$ are integrally dependent on $\eta_{k+1}, \eta_{k+2}, \cdots, \eta_n$, q.e.d.

(c) *If $\omega_0, \omega_1, \cdots, \omega_r$ are forms in $\eta_0, \eta_1, \cdots, \eta_n$, all of the same degree h, and if the ideal $(\omega_0, \omega_1, \cdots, \omega_r)$ is irrelevant, then the η's are integrally dependent on the ω's.* For if $\eta_0^{(h)}, \eta_1^{(h)}, \cdots$ form a linear base for the forms of degree h in the η's and if we include the ω's in this base, then applying (b) to the ring $K[\eta_0^{(h)}, \eta_1^{(h)}, \cdots]$, we conclude that $\eta_0^h, \eta_1^h, \cdots, \eta_n^h$ are integrally dependent on the ω's.

(d) We obviously can select (in many ways) $r+1$ forms $\zeta_0, \zeta_1, \cdots, \zeta_r$ in $\mathfrak{o}^*$ such that the ideal $(\zeta_0, \zeta_1, \cdots, \zeta_r)$ be irrelevant. We can then find exponents σ_i such that the forms $\omega_i = \zeta_i^{\sigma_i}$ be of like degree. Then it follows, by (c), that the η's are integrally dependent on $\omega_0, \omega_1, \cdots, \omega_r$. This completes the proof of the extended "normalization theorem" for homogeneous integral domains.

From an homogeneous integral domain $K[\eta_0, \eta_1, \cdots, \eta_n]$, of degree of transcendency $r+1$, we get an arbitrary integral domain $K[\xi_1, \xi_2, \cdots, \xi_n]$, of degree of transcendency r, by putting $\xi_i = \eta_i/\eta_0$. We apply step (d) above, and we observe that of the $r+1$ forms ζ_i, one, say ζ_0, can be taken arbitrarily. If we put $\zeta_0 \doteq \eta_0$, then ω_0 is a power of η_0, say $\omega_0 = \eta_0^h$. It is then seen immediately that the r elements $\xi_i' = \omega_i/\eta_0^h$, $i = 1, 2, \cdots, r$, are polynomials in the ξ's and that $\xi_1, \xi_2, \cdots, \xi_n$ are integrally dependent on $\xi_1', \xi_2', \cdots, \xi_r'$. This completes the proof.

Let $(\eta_0, \eta_1, \cdots, \eta_n)$ be the general point of V (the coordinates η_i are homogeneous[20]) and let $P = \mathrm{K}[\eta_0, \eta_1, \cdots, \eta_n]$. Let $\overline{P}$ be integral closure of P in its quotient field. We first need to establish *in the most general case* that $\overline{P}$ is a finite P-module. By the normalization theorem of Emmy Noether, let $\zeta_0, \zeta_1, \cdots, \zeta_r$ be $r+1$ algebraically independent elements in P such that every element of P is integrally dependent on $\mathrm{K}[\zeta_0, \zeta_1, \cdots, \zeta_r]$. This last ring is a polynomial ring and we shall denote it by R. Since the field $\Sigma = \mathrm{K}(\eta_0, \eta_1, \cdots, \eta_n)$ is a finite algebraic extension of the quotient field of *the polynomial ring R*, it follows that *the integral closure $\overline{R}$ of R in Σ is a finite R-module.* This result, for arbitrary ground fields, has been proved by F. K. Schmidt [5]. Since $P \subseteq \overline{R}$, it follows that every element of $\overline{P}$ is integrally dependent on $\overline{R}$. But since $\overline{R}$ is a finite R-module and is therefore a chain-theorem ring, it is well known that every element of Σ which is integrally dependent on $\overline{R}$ is also integrally dependent on R, that is, belongs to $\overline{R}$. Hence $\overline{P} \subseteq \overline{R}$, that is, $\overline{P} = \overline{R}$. Thus $\overline{P}$ is a finite R-module, q.e.d.

(20) To avoid repetitions, we stipulate from now on that whenever the subscript in a set of coordinates begins with 0, the coordinates are homogeneous.

Since $\bar{P}$ is a finite P-module, it is a finite integral domain over K. Let $\bar{P}=K[\bar{\zeta}_0, \bar{\zeta}_1, \cdots, \bar{\zeta}_m]$. As in the quoted paper [7] (see p. 290), we may assume also here that the $\bar{\zeta}_i$ are homogeneous elements.

Then by exactly the same procedure as that carried out in our quoted paper we can show that if $\omega_0^*, \omega_1^*, \cdots, \omega_\mu^*$ is a linear K-basis for all the homogeneous elements of $\bar{P}$ of a given degree δ, then for suitable integers δ it will be true that every homogeneous element in $\bar{P}$ whose degree is a multiple $\rho\delta$ of δ, $\rho>0$, is necessarily a form of degree ρ in $\omega_0^*, \omega_1^*, \cdots, \omega_\mu^*$. From that we concluded in the quoted paper that for such an integer δ *the ring* $P^*=K[\omega_0^*, \omega_1^*, \cdots, \omega_\mu^*]$ *is integrally closed in its quotient field*, whence the variety $\bar{V}$ whose general point is $(\omega_0^*, \omega_1^*, \cdots, \omega_\mu^*)$ is normal. This variety $\bar{V}$ we termed a derived normal variety of V. It was pointed out to me by Irvin Cohen that the above conclusion fails to hold true if K is not maximally algebraic in the field Σ of rational functions on V, that is, *if V is not absolutely irreducible* (see [10, Lemma 4, p. 64]). For in this case the elements of Σ which are algebraic over K but are not in K, that is, *the homogeneous elements of degree zero*, are certainly not in the ring P^*. However, it is still true that P^* contains all homogeneous integral quantities (that is, the homogeneous elements of the quotient field of P^* which are integrally dependent of P^*) *of positive degree*. Hence if α^* is any integral quantity in the quotient field of P^*, then the products $\alpha^*\omega_i^*$, $i=0, 1, \cdots, \mu$, belong to P^*, since they are sums of homogeneous integral quantities of positive degree. It follows that *the irrelevant ideal* $P^*(\omega_0^*, \omega_1^*, \cdots, \omega_\mu^*)$ *is the conductor of the ring* P^* *with respect to its integral closure in the quotient field of* P^*. This implies that the variety $\bar{V}$ is *locally normal* in the sense of Definition 3 given later on in this section (see also [7, Theorem 13, p. 286]). For our purpose a locally normal variety is just as effective as a normal variety. We shall continue to call $\bar{V}$ the derived normal model of V, it being understood that if V is not absolutely irreducible then $\bar{V}$ is only locally normal. It may be well to point out at this stage the self-evident fact that if V is not absolutely irreducible, the field Σ does not possess at all normal models *over* K.

For the general theory of birational correspondences it is necessary to establish first some properties of the birational correspondence between a given model V and a derived normal model $\bar{V}$ of V. For it will follow from the properties of this birational correspondence that the properties of a birational correspondence between any two models V, V' can be readily deduced from the properties of the birational correspondence between the derived normal models of V and V'. Therefore, there is no loss of generality if the theory is restricted to normal models. On the other hand, the emphasis on normal models is advantageous, both from a technical and a conceptual standpoint, since in the case of normal varieties the theory of birational correspondences is free from many accidental complications and irrelevant exceptions which one often encounters on non-normal varieties.

Let us therefore consider the case in which one of the two birationally equivalent varieties V, V' is a derived normal model of the other. Let, say, V' be a derived normal model of V and let $(\eta_0, \eta_1, \cdots, \eta_n)$ and $(\eta_0', \eta_1', \cdots, \eta_m')$ be the general points of V and of V', respectively. Let h be the degree of homogeneity of V'. The elements η_i' form then a linear base for the elements of the field $K(\eta_0, \eta_1, \cdots, \eta_n)$ which are integrally dependent on $\eta_0, \eta_1, \cdots, \eta_n$ and which are homogeneous of degree h. Moreover, the ring $P' = K[\eta_0', \eta_1', \cdots, \eta_m']$ is integrally closed (in its quotient field), or at any rate contains all homogeneous integers of positive degree.

It will be convenient to use an auxiliary projective model V^* defined as follows. Let $\eta_0^*, \eta_1^*, \cdots, \eta_s^*$ be a linear base for the forms of degree h in $\eta_0, \eta_1, \cdots, \eta_n$, with coefficients in K. We take as V^* the variety whose general point is $(\eta_0^*, \eta_1^*, \cdots, \eta_s^*)$.

LEMMA 5. *The birational correspondence between V and V^* is* (1, 1) *without exceptions*[21]. *Any two corresponding irreducible subvarieties of V and V' have the same dimension and the same quotient ring.*

Proof. Let W and W^* be two corresponding irreducible subvarieties of V and V^*, respectively. We assume that $\eta_0 \neq 0$ on W. Since the η^*'s constitute a linear base for the forms of degree h in the η's, we may assume that η_0^h is one of the η^*'s, say $\eta_0^h = \eta_0^*$. If v be a valuation of center W and W^*, then $v(\eta_i/\eta_0) \geqq 0$, since $\eta_0 \neq 0$ on W. From this it follows that $v(\eta_j^*/\eta_0^*) \geqq 0$, for $j = 0, 1, \cdots, s$, and hence (see I.3) $\eta_0^* \neq 0$ *on* W^*. Now let ζ be any element of $Q(W)$, say $\zeta = \phi(\eta)/\psi(\eta)$, where ϕ and ψ are forms of like degree and where $\psi(\eta) \neq 0$ on W. We may assume that the common degree of ϕ and ψ is a multiple of h, say ρh, since we can multiply both ϕ and ψ by any power of η_0 without destroying the inequality: $\psi \neq 0$ on W. But if ϕ and ψ are of degree ρh, then they can be expressed as forms of degree ρ in the η^*'s: $\phi(\eta) = \phi^*(\eta^*)$, $\psi(\eta) = \psi^*(\eta^*)$. Since $\psi(\eta) \neq 0$ on W, we have $v(\psi(\eta)/\eta_0^{\rho h}) = 0$. Hence $v(\psi^*(\eta^*)/\eta_0^{*\rho}) = 0$, and this shows that $\psi^*(\eta^*) \neq 0$ on W^*. Since $\zeta = \phi^*(\eta^*)/\psi^*(\eta^*)$, *we conclude that* $\zeta \in Q(W^*)$.

A quite similar argument shows that if $\zeta \in Q(W^*)$, then $\zeta \in Q(W)$. Hence the quotient rings $Q(W)$ and $Q(W^*)$ coincide, and from this our lemma follows in view of Theorem 3 (I.3).

The lemma shows that as far as the study of the birational correspondence between V and V' is concerned, it is permissible to replace V by V^*. Let us see therefore how V' is related to V^*.

Every homogeneous element in $K(\eta_0, \eta_1, \cdots, \eta_n)$, of degree h,—and in particular each element η_i'—can be written as a quotient of two forms in the η's whose degrees are multiples of h. Any such quotient is a quotient of

(21) When we say that a birational correspondence between two varieties V and V' is (1, 1) without exceptions, we mean that it is (1, 1) as a correspondence between the irreducible subvarieties of V and the irreducible subvarieties of V'.

two forms in the η^*'s. Hence $\eta_i' \in K(\eta_0^*, \eta_1^*, \cdots, \eta_s^*)$. Conversely, each element η_i^*, being homogeneous of degree h, is a linear form in the η''s. We therefore conclude that *the two fields* $K(\eta_0^*, \eta_1^*, \cdots, \eta_s^*)$ and $K(\eta_0', \eta_1', \cdots, \eta_m')$ *coincide.*

The elements η_i^*, which as elements of the field $K(\eta_0, \eta_1, \cdots, \eta_n)$ are homogeneous of degree h, as elements of the field $K(\eta_0^*, \eta_1^*, \cdots, \eta_s^*)$ are to be regarded as homogeneous, of degree 1. The same remark applies to the elements η_i'. Moreover, the elements η_i' constitute a linear base for the elements of the field $K(\eta_0^*, \eta_1^*, \cdots, \eta_s^*)$ which are homogeneous of degree 1 and which are integrally dependent on $\eta_0^*, \eta_1^*, \cdots, \eta_s^*$. We conclude from all this that *V' is also a derived normal variety of V^*, of degree of homogeneity* 1. Thus, while not every variety V possesses a derived normal variety of degree of homogeneity 1, we may nevertheless assume—and we do so assume—that we had originally $h=1$; this assumption amounts to replacing V by V^*.

Now that we have $h=1$, it follows that η's are linear combinations of the η''s, whence *V is a projection of the normal model V'*. Moreover, the ring $P'=K[\eta']$ is now either the integral closure of the ring $P=K[\eta]$ in its quotient field or contains at any rate all homogeneous integral quantities of positive degree.

Let now W and W' be corresponding irreducible subvarieties of V and V', respectively. We assume that $\eta_0 \neq 0$ on W and that $\eta_0' = \eta_0$. *I assert that $\eta_0' \neq 0$ on W'.* To see this we have only to show that $v(\eta_i'/\eta_0') \geqq 0$, $i=0, 1, \cdots, m$, for at least one valuation of center W'. We take as v a valuation which has also W as center on V. We write the relation of integral dependence for η_i' over P. It is of the form (see our paper [7, p. 286, equation (33)]):

$$\eta_i'^{\nu} + a_1(\eta)\eta_i'^{\nu-1} + \cdots + a_\nu(\eta) = 0,$$

where $a_j(\eta)$ is a form of degree j in $\eta_0, \eta_1, \cdots, \eta_n$. If we divide this equation by $\eta_0'^{\nu}$, we see that the quotients η_i'/η_0' are integrally dependent on the quotients $\eta_1/\eta_0, \eta_2/\eta_0, \cdots, \eta_n/\eta_0$. Since $\eta_0 \neq 0$ on W, these quotients are in R_v, and this proves our assertion.

Let $\mathfrak{P}$ and $\mathfrak{P}'$ be the prime homogeneous ideals of W and of W' in the ring P and P', respectively. From the fact that $\eta_0 \neq 0$ on W and $\eta_0' (=\eta_0) \neq 0$ on W' and from the very definition of the center of a valuation, it follows immediately that $\mathfrak{P} = \mathfrak{P}' \cap P$. Since the elements of P' are integrally dependent on P, there is only a finite number of prime ideals in P' which contract to $\mathfrak{P}$. These ideals are all homogeneous[22] and of the same dimension as $\mathfrak{P}$. We therefore reach the following conclusion:

To each irreducible subvariety W' of V' there corresponds a unique subvariety W to V, while to each irreducible subvariety W of V there corresponds a finite

[22] The prime ideals $\mathfrak{P}'$ which contract to $\mathfrak{P}$ are the minimal primes of the extended ideal $P' \cdot \mathfrak{P}$ and therefore are homogeneous, by Theorem 2 (I.1).

number of subvarieties W' of V'. Two corresponding varieties W and W' have the same dimension.

We now investigate the relationship between the quotient rings of two corresponding varieties W and W'. We pass to the rings $\mathfrak{o}=K[\xi_1, \xi_2, \cdots, \xi_n]$ and $\mathfrak{o}'=K[\xi_1', \xi_2', \cdots, \xi_m']$ of the nonhomogeneous coordinates $\xi_i=\eta_i/\eta_0$ and $\xi_i'=\eta_i'/\eta_0'$, where $\eta_0=\eta_0'$. Here $\mathfrak{o}'$ is the integral closure of $\mathfrak{o}$. Let $\mathfrak{p}$ be the prime $\mathfrak{o}$-ideal of W. Any W' which corresponds to W will be at finite distance with respect to the coordinates ξ_i' and will be given in $\mathfrak{o}'$ by a prime ideal which contracts to $\mathfrak{p}$. Let $\mathfrak{p}_1', \mathfrak{p}_2', \cdots, \mathfrak{p}_\nu'$ be the prime $\mathfrak{o}'$-ideals which contract to $\mathfrak{p}$, and let $W_1', W_2', \cdots, W_\nu'$ be the corresponding subvarieties of V'. Let

$$\mathfrak{J} = Q(W) = \mathfrak{o}_{\mathfrak{p}}, \qquad \mathfrak{J}_i' = Q(W_i') = \mathfrak{o}'_{\mathfrak{p}_i'},$$

and let $\mathfrak{m}$ and $\mathfrak{m}_i'$ denote the ideals of non-units in $\mathfrak{J}$ and $\mathfrak{J}_i'$, respectively. We have $\mathfrak{J}_i'\supseteq\mathfrak{J}$ and $\mathfrak{J}_i'$ is integrally closed. Let $\mathfrak{J}^*$ denote the integral closure of $\mathfrak{J}$. Then $\mathfrak{J}_i'\supseteq\mathfrak{J}^*\supset\mathfrak{o}'$ and we can consider the ideals $\mathfrak{m}_i^*=\mathfrak{m}_i'\cap\mathfrak{J}^*$. *The ideals* $\mathfrak{m}_1^*, \mathfrak{m}_2^*, \cdots, \mathfrak{m}_\nu^*$ *are distinct*, since $\mathfrak{m}_i^*\cap\mathfrak{o}'=\mathfrak{m}_i'\cap\mathfrak{o}'=\mathfrak{p}_i'$. *The ideals* $\mathfrak{m}_i^*$ *contract to one and the same ideal in* $\mathfrak{J}$, *namely to* $\mathfrak{m}$, since $\mathfrak{p}_i'\cap\mathfrak{o}=\mathfrak{p}$ and $\mathfrak{m}=\mathfrak{J}\cdot\mathfrak{p}$. The quotient ring of $\mathfrak{m}_i^*$ in $\mathfrak{J}^*$ is contained in $\mathfrak{J}_i'$ since $\mathfrak{m}_i^*=\mathfrak{m}_i'\cap\mathfrak{J}^*$. On the other hand, the quotient ring $\mathfrak{J}_i'$ is contained in the quotient ring of $\mathfrak{m}_i^*$, since $\mathfrak{m}_i^*\cap\mathfrak{o}'=\mathfrak{p}_i'$. *Hence the quotient ring of* $\mathfrak{m}_i^*$ *in the ring* $\mathfrak{J}^*$ *coincides with* $\mathfrak{J}_i'$.

The foregoing properties of $\mathfrak{J}^*$ can also be derived from the general theory of quotient rings (see I.2 ([1])). It is only necessary to observe that $\mathfrak{J}^*$ coincides with the quotient ring $\mathfrak{o}_S'$, where $S=\mathfrak{o}-\mathfrak{p}$. From this remark it follows immediately that the ideals $\mathfrak{m}_1^*, \mathfrak{m}_2^*, \cdots, \mathfrak{m}_\nu^*$ are the only prime ideals of $\mathfrak{J}^*$ which contract to $\mathfrak{m}$. The connection between the quotient ring $Q(W)$ and the ν quotient rings $Q(W_i')$ is therefore fully established. Reassuming, we can now state the following theorem:

THEOREM 7. *The birational correspondence between an irreducible algebraic variety V and a derived normal variety V' of V has the following properties:*

(A) *Two corresponding subvarieties W and W' of V and V', respectively, have the same dimension, and we have: $Q(W)\subseteq Q(W')$.*

(B) *Given W', the corresponding W is uniquely determined, while to a given W there corresponds a finite number of varieties W'.*

(C) *If $\mathfrak{J}^*$ denotes the integral closure of the quotient ring $\mathfrak{J}=Q(W)$, and if $\mathfrak{m}_1^*, \mathfrak{m}_2^*, \cdots, \mathfrak{m}_\nu^*$ are the prime ideals in $\mathfrak{J}^*$ which contract to the ideal of non-units in $\mathfrak{J}$, then there are exactly ν varieties $W_1', W_2', \cdots, W_\nu'$ which correspond to W, and for a suitable ordering of the indices we will have $Q(W_i')=\mathfrak{J}^*_{\mathfrak{m}_i^*}$.*

COROLLARY 1. *The birational correspondence between any two derived normal varieties of V is* (1, 1) *without exceptions, and corresponding subvarieties have the same quotient ring.*

Other corollaries follow from Theorem 7. We first give the following definitions:

DEFINITION 2. *If $Q(W)$ is integrally closed, then V is said to be locally normal at W.*

DEFINITION 3. *If V is locally normal at each of its irreducible subvarieties, then V is said to be a locally normal variety*(23).

A locally normal variety is characterized by the property that it is normal in the affine space for every choice of the nonhomogeneous coordinates. For that it is necessary and sufficient (see our paper [7, Theorem 13, p. 286]) that the conductor of the ring $P=K[\eta_0, \eta_1, \cdots, \eta_n]$ with respect to the integral closure of P be an irrelevant ideal.

COROLLARY 2. *If V is locally normal at W then to W there corresponds a unique subvariety W' of the derived normal model V', and the quotient rings $Q(W)$, $Q(W')$ coincide.*

COROLLARY 3. *If V is locally normal, then the birational correspondence between V and a derived normal variety V' of V is* (1, 1) *without exceptions, and the correspondence preserves quotient rings.*

COROLLARY 4. *The irreducible subvarieties of V to which there corresponds more than one variety on V' all lie on the subvariety C of V which is defined by the conductor $\mathfrak{C}$ of the ring $P=K[\eta_0, \eta_1, \cdots, \eta_n]$ with respect to the integral closure of P. Outside of C, the birational correspondence between V and V' is* (1, 1) *and it preserves quotient rings.*

3. **The fundamental elements of a birational correspondence.** We consider a birational correspondence T between two *locally normal* varieties V and V' (the general case is discussed briefly at the end of this section). Let W be an irreducible subvariety of V.

DEFINITION 4. *We say that W is* (1) *regular,* (2) *irregular, or* (3) *fundamental for T if there exists a W' on V' such that $W'=T(W)$ and, respectively,* (1) $Q(W)=Q(W')$, (2) $Q(W)\supset Q(W')$ or (3) $Q(W)\not\supseteq Q(W')$.

The following theorem is merely a statement of some properties of regular, irregular and fundamental varieties which follow directly from the definition and from Theorem 6 (II.1) and which we shall use very frequently:

THEOREM 8.

(A) *If W is regular or irregular, then to W there corresponds a unique W' on V'.*

(23) For V to be locally normal it is sufficient that V be locally normal at each of its points. For if W is any irreducible subvariety of V and if P is any point of W, then $Q(W)$ is also the quotient ring of a prime ideal in $Q(P)$. If $Q(P)$ is integrally closed, also $Q(W)$ is integrally closed.

(B) *If W is regular for T and if $W'=T(W)$, then W' is of the same dimension as W and is regular for T^{-1}.*

(C) *If W is irregular and $W'=T(W)$, then W' is fundamental for T^{-1}, and the dimension of W' is less than or equal to the dimension of W.*

(D) *If W is fundamental, then $Q(W)\not\supseteq Q(W')$ for any W' which corresponds to W.*

(E) *A necessary and sufficient condition that W not be fundamental is that for a suitable choice of the nonhomogeneous coordinates $\xi_1', \xi_2', \cdots, \xi_m'$ of the general point of V', the ring $\mathfrak{o}'$ of these coordinates be contained in $Q(W)$.*

The proof of (E) is immediate. For assume that $\mathfrak{o}'\subseteq Q(W)$ and let $\mathfrak{p}'$ be the prime $\mathfrak{o}'$-ideal which is the contraction of the ideal of non-units of $Q(W)$. Then if W' is the irreducible subvariety of V' which is defined by $\mathfrak{p}'$, then $Q(W')\subseteq Q(W)$ and by Theorem 3 every valuation of center W on V has center W' on V'.

The sets of regular, irregular and fundamental varieties are mutually exclusive. We now give the following definition:

Definition 5. *A birational correspondence T is regular, if every W is regular for T.*

It is clear that if T is regular, then also T^{-1} is regular (Theorem 8 (B)) A regular birational correspondence is a (1, 1) correspondence, without exceptions, and it preserves quotient rings. We have encountered examples of regular birational correspondences in the preceding section (Lemma 5; Theorem 7, Corollaries 1 and 3).

The next theorem shows that Theorem 8 (A) expresses a *characteristic property* of a non-fundamental W.

Theorem 9. *If to W there corresponds a unique subvariety W' of V', then W is not fundamental for T (hence is either regular or irregular).*

Proof. Since V is locally normal, the quotient ring $Q(W)$ is integrally closed. By the fundamental theorem on principal orders[14], $Q(W)$ is the intersection of the valuation rings which contain $Q(W)$. Let R_{v_1} be one of these valuation rings. Since $R_{v_1}\supseteq Q(W)$, it follows that the center of v_1 on V is either W or a subvariety W_1 of V which properly contains W (Theorem 3, I.3). In the second case there exists a valuation v which is compounded with v_1 and has center W (Lemma 4, I.4). Since $R_v\subset R_{v_1}$ we can omit R_{v_1} from the set of valuation rings which contain $Q(W)$, without affecting the intersection of these rings. *Hence $Q(W)$ is also the intersection of the valuation rings which belong to valuations of center W.* Since, by hypothesis, all valuations of center W on V have the same center W' on V', it follows that the corresponding valuation rings all contain $Q(W')$. Consequently $Q(W)\supseteq Q(W')$, as was asserted.

COROLLARY. *The dimension of a fundamental variety W cannot exceed $r-2$.*

For if W has dimension $r-1$, then $Q(W)$ is itself a valuation ring, namely the valuation ring of a divisor v. Therefore W is the center of only one valuation, namely of the divisor v.

THEOREM 10. *If W is fundamental, then to W there correspond on V' infinitely many varieties W'.*

Proof. We shall prove that if to W there corresponds on V' only a finite number of varieties, then W is not fundamental.

Let $W_1', W_2', \cdots, W_h'$ be the irreducible subvarieties of V' which correspond to W. If v is any valuation of center W, then R_v must contain at least one of the h quotient rings $Q(W_i')$. Hence R_v contains the intersection of these quotient rings. Since $Q(W)$ is the intersection of all R_v, it follows that $Q(W)$ *contains the intersection of the h quotient rings $Q(W_i')$.*

We can find a form $\phi(\eta_0', \eta_1', \cdots, \eta_m')$, of a sufficiently high degree ν, such that $\phi \neq 0$ on W_i', $i=1, 2, \cdots, h$. We pass from V' to the variety V_1' whose general point is defined by a linear K-basis of the forms of degree ν in $\eta_0', \eta_1', \cdots, \eta_m'$. By Lemma 5, V' and V_1' are in regular birational correspondence, hence we may replace in our proof V' by V_1'. We may therefore assume that ϕ is one of the elements η_i', say $\phi = \eta_0'$. From the fact that $\eta_0' \neq 0$ on W_i', $i=1, 2, \cdots, h$, it follows that the ring $\mathfrak{o}'$ of the nonhomogeneous coordinates $\xi_i' = \eta_i'/\eta_0'$ is contained in each quotient ring $Q(W_i')$. Since the intersection of the rings $Q(W_i')$ is contained in $Q(W)$, our theorem follows from Theorem 8 (E).

We shall now discuss briefly the general case in which V and V' are not locally normal. Let $\overline{V}$ and $\overline{V}'$ be derived normal varieties of V and V' respectively. Let W be an irreducible subvariety of V and let $\overline{W}_1, \overline{W}_2, \cdots, \overline{W}_h$ be the irreducible varieties on $\overline{V}$ which correspond to W (Theorem 7 (B)). We shall denote by $\overline{T}$ the birational correspondence between $\overline{V}$ and $\overline{V}'$.

DEFINITION 6. *The variety W is regular for T, if each $\overline{W}_i$, $i=1, 2, \cdots, h$, is regular for $\overline{T}$; W is fundamental for T, if at least one of the varieties $\overline{W}_i$ is fundamental for $\overline{T}$; W is irregular for T, if it is neither regular nor fundamental, that is, if no $\overline{W}_i$ is fundamental for $\overline{T}$ and if at least one $\overline{W}_i$ is irregular for $\overline{T}$.*

Of the theorems proved in this section for locally normal varieties, Theorems 9 and 10 continue to hold in the general case. The validity of Theorem 10 is obvious. As to Theorem 9, the proof is as follows. If T is single-valued at W, say $T(W) = W'$, then any irreducible subvariety of $\overline{V}'$ which corresponds to $\overline{W}_i$ under $\overline{T}$ ($i=1, 2, \cdots, h$) must be among the irreducible subvarieties of $\overline{V}'$ which correspond to W' in the birational correspondence between V' and $\overline{V}'$. Hence to each $\overline{W}_i$ there can correspond on $\overline{V}'$ only a finite number of varieties. Therefore no $\overline{W}_i$ is fundamental (Theorem 10) for $\overline{T}$, and therefore, by definition, W is not fundamental for T.

In particular, if $T(W)=W'$ and if $Q(W)\supseteq Q(W')$, then W is not fundamental. This follows from Theorem 6, II.1.

On the other hand, other results established for locally normal varieties do not generalize to varieties which are not locally normal. For instance, parts (A) and (B) of Theorem 8 cease to be true in the general case. Also the defining property of a non-fundamental variety used in Definition 4 ceases to be a property of non-fundamental varieties in the general case, that is, if W is not fundamental that does not mean that there must exist a W' such that $W'=T(W)$ and $Q(W)\supseteq Q(W')$. Also the condition stated in Theorem 8 (E) is sufficient, but no longer necessary.

Note that according to Definition 6 the birational correspondence between a variety V and derived normal variety of V is free from fundamental elements on either variety.

We shall agree to use Definition 5 of regular birational correspondences also in the case of varieties which are not locally normal.

4. **A question of terminology.** At this stage it becomes necessary to point out and to discuss the difference between our terminology and the terminology used heretofore in the literature. This difference concerns the meaning of the term "fundamental" and our use of the new term "irregular."

In the case of algebraic surfaces it is the sense of the old terminology that both points *and curves* can be fundamental: a point P is fundamental if it is transformed into a curve Γ', and any such curve Γ', which is then the transform of a point, is "fundamental." As far as the notion of a fundamental point is concerned this is in agreement with our terminology, from Theorem 8 (A) and Theorem 10. However, by Theorem 9, corollary, a curve on an algebraic surface can never be fundamental in our sense. The "fundamental" curves in the sense of the old terminology are irregular curves in our sense.

The reasons for our terminology—or better—the inadequacy of the old terminology([24]) become apparent in the case of higher varieties. Let us consider, for instance, a birational correspondence T between two 3-dimensional varieties V and V'. Again, according to the old terminology we may have "fundamental" loci of all dimensions from 0 to 2. As far as fundamental points and "fundamental" surfaces are concerned, the situation is the same as in the case of algebraic surfaces: there is complete agreement on fundamental points, while according to our terminology there are definitely no "fundamental" surfaces, but only irregular surfaces. It is, however, the use of the term "fundamental curve" that brings out some significant facts.

I can find no clear-cut definition of a fundamental curve in the literature. This much is certain: if a curve Γ is such that $T(\Gamma)$ is a surface, or if Γ corresponds to each point of another curve, then in the old terminology (and

([24]) The best justification for our terminology is its own logical consistency. We call fundamental a variety W if and only if the birational correspondence T is infinitely many-valued at W. Otherwise W is either regular or irregular.

also in our terminology) Γ is fundamental (respectively, of the "first" or of the "second kind"). Suppose, however, that the transform of Γ is a single point. I am not certain whether or not such a curve is fundamental in the sense of the old terminology. If it is, then the terminology is confusing, since we are dealing here with a curve at which the birational transformation is single-valued. If it is not, then the terminology is inconsistent, in view of the use of the term "fundamental" surface, since in both cases we are dealing with a W such that $T(W)$ is ***unique*** and is of ***lower*** dimension than W.

It is quite possible that in the old terminology no special name has ever been assigned to a curve Γ such that $T(\Gamma)$ is a single point. If that is the case, then this is probably due to the fact that, as a rule, only nonsingular models have been considered in the literature. If the three-dimensional varieties V and V' are nonsingular, then a curve Γ which is transformed into a point necessarily lies on a surface which is transformed into a curve (see II.10, Theorem 17, corollary). Thus, such a curve Γ always lies on a "fundamental" surface, and there seemed to be no compelling reason for giving these curves a special name. However, in the case of singular models it may very well happen that a curve Γ whose dimension is ***lowered*** by the birational transformation T and at which T is ***single-valued*** (these two properties imply that W is irregular; see Theorem 8 (B) and Theorem 9) does not lie on any surface having the same properties (compare with Theorem 17, II.10). Some term for such a curve is necessary, and the term "fundamental" we reject for reasons given above.

5. **The join of two birationally equivalent varieties.** Let $(\eta_0, \eta_1, \cdots, \eta_n)$ and $(\eta_0', \eta_1', \cdots, \eta_m')$ be the general points, respectively, of V and of V', where V and V' are our two birationally equivalent varieties. Since the quotients η_i'/η_0' are rational functions of the quotients η_j/η_0, the η''s are proportional to forms of like degree in the η's:

$$\begin{aligned}\eta_0' : \eta_1' : \cdots : \eta_m' \\ = \phi_0(\eta_0, \eta_1, \cdots, \eta_n) : \phi_1(\eta_0, \eta_1, \cdots, \eta_n) : \cdots : \phi_m(\eta_0, \eta_1, \cdots, \eta_n).\end{aligned} \tag{4}$$

DEFINITION 7. *The irreducible algebraic variety V^* whose general point has the $(n+1)(m+1)$ products $\eta_i\phi_j$ as homogeneous coordinates is called the join of V and V'.*

We shall denote the products $\eta_i\phi_j$ by η_{ij} and the quotients η_i/η_0, η_j'/η_0' and η_{ij}/η_{00} by ξ_i, ξ_j' and ξ_{ij}, respectively. We have then:

$$\xi_{i0} = \xi_i, \qquad \xi_{0j} = \xi_j', \qquad \xi_{ij} = \xi_i\xi_j', \qquad i, j \neq 0, \tag{5}$$

and from these relations it follows that V^* is birationally equivalent to V (and to V'). Moreover, if we take as nonhomogeneous coordinates of the general point of V^* the quotients of the η_{ij}'s by a fixed η_{ij}, say by η_{00}, then the ring of these coordinates is, by (5), the join of the two rings of nonhomogene-

ous coordinates relative to V and V'. In symbols: if

$$\mathfrak{o} = K[\xi_1, \xi_2, \cdots, \xi_n], \qquad \mathfrak{o}' = K[\xi_1', \xi_2', \cdots, \xi_m'],$$
$$\mathfrak{o}^* = K[\xi_{10}, \xi_{20}, \cdots, \xi_{nm}],$$

then $\mathfrak{o}^* = (\mathfrak{o}, \mathfrak{o}')$.

THEOREM 11. *In the birational correspondence T^* between V and V^* there corresponds to any irreducible subvariety W^* of V^* a unique subvariety W of V. If a given W on V is not fundamental for T, then it is regular for T^*. Similarly for V' and V^*.*

Proof. We may assume that W^* is at finite distance with respect to the nonhomogeneous coordinates ξ_{ij}. If v is any valuation of center W^*, then $R_v \supseteq Q(W^*) \supset \mathfrak{o}^* \supseteq \mathfrak{o}$, whence the center W of v on V is at finite distance with respect to the nonhomogeneous coordinates ξ_i. Similarly for V', W' and the ξ_j'. But then, if $\mathfrak{p}^*$ is the prime $\mathfrak{o}^*$-ideal of W^*, the prime ideals $\mathfrak{p}$ and $\mathfrak{p}'$ of W and W', in the rings $\mathfrak{o}$ and $\mathfrak{o}^*$, respectively, are necessarily the contracted ideals of $\mathfrak{p}^*$, that is, $\mathfrak{p} = \mathfrak{p}^* \cap \mathfrak{o}$, $\mathfrak{p}' = \mathfrak{p}^* \cap \mathfrak{o}'$, and consequently W and W' are uniquely determined by W^*. Notice that the quotient rings $Q(W)$ and $Q(W')$ are subrings of $Q(W^*)$:

$$(6) \qquad Q(W) \subseteq Q(W^*), \qquad Q(W') \subseteq Q(W^*).$$

To prove the second part of the theorem, let $T(W) = W'$ and let us assume that W and W' are at finite distance with respect to the nonhomogeneous coordinates ξ_i and ξ_j'. Let v be any valuation of center W and W' on V and V', respectively. The center W^* of v on V^* will be at finite distance with respect to the coordinates ξ_{ij}, since $R_v \supset \mathfrak{o}$, $R_v \supset \mathfrak{o}'$ and therefore $R_v \supset \mathfrak{o}^*$.

Let us first consider the case in which V is locally normal. Since W is not fundamental for T, we have $Q(W) \supseteq Q(W') \supset \mathfrak{o}'$. Therefore $Q(W) \supset \mathfrak{o}^*$, that is, $\mathfrak{o}_\mathfrak{p} \supset \mathfrak{o}^*$. Since $\mathfrak{p}^* \cap \mathfrak{o} = \mathfrak{p}$, it follows that $\mathfrak{o}_\mathfrak{p}$ contains the ring $\mathfrak{o}^*_{\mathfrak{p}^*}$, that is, $Q(W) \supseteq Q(W^*)$. Hence, by (6), $Q(W) = Q(W^*)$, whence W is regular for T^*.

To prove the theorem in the general case we first observe that if $\bar{V}$ and $\bar{V}'$ denote derived normal varieties of V and of V', respectively, then any derived normal variety $\bar{V}^*$ of the join of V and V' is in regular birational correspondence with any derived normal variety of the join of $\bar{V}$ and $\bar{V}'$. The proof is straightforward and consists in the obvious remark that the integral closure of the ring $(\mathfrak{o}, \mathfrak{o}')$ is the same as integral closure of the ring $(\mathfrak{o}_1, \mathfrak{o}_1')$, where $\mathfrak{o}_1$ and $\mathfrak{o}_1'$ are the integral closures of $\mathfrak{o}$ and of $\mathfrak{o}'$, respectively. Now let $\bar{W}$ be any of the irreducible subvarieties of $\bar{V}$ which correspond to W. To prove that W is regular for T^* we have only to prove (Definition 6, II.3) that $\bar{W}$ is regular for the birational correspondence between $\bar{V}$ and $\bar{V}^*$. Since W, by hypothesis, is not fundamental for T, it follows (Definition 6) that $\bar{W}$ is not fundamental for the birational correspondence between $\bar{V}$ and $\bar{V}'$. Since $\bar{V}$ is locally normal, it follows by the case just considered that $\bar{W}$ is regular

for the birational correspondence between $\bar{V}$ and the join of $\bar{V}$ and $\bar{V}'$. Since a derived normal variety of this join is, by the remark made above, in regular birational correspondence with $\bar{V}^*$, it follows that $\bar{W}$ is regular for the birational correspondence between $\bar{V}$ and $\bar{V}^*$, as was asserted.

COROLLARY. *If P and P' are corresponding points of V and V', then there is only a finite number of points P^* on V^* which correspond to both P and P'. If V (or V') is locally normal, then the number of such points P^* can be greater than 1 only if P (or P') is a fundamental point of T (or of T^{-1}).*

For if we identify the varieties W, W' and W^* of the preceding proof with the points P, P' and P^*, respectively, we see that $\mathfrak{p}^*$ must be a zero-dimensional prime divisor of the ideal $\mathfrak{o}^*\cdot(\mathfrak{p}, \mathfrak{p}')$. Since this ideal is pure zero-dimensional, the number of possible ideals $\mathfrak{p}^*$ is finite. The second half of the corollary follows directly from the second half of the preceding theorem.

If the ground field K is algebraically closed, the ideal $\mathfrak{o}^*\cdot(\mathfrak{p}, \mathfrak{p}')$ is itself prime, whenever $\mathfrak{p}$ and $\mathfrak{p}'$ are both zero-dimensional. Hence if K is algebraically closed, then not only does every point P^* of V^* determine uniquely a pair of *corresponding* points P, P' of V and V', respectively, but, conversely, every such pair determines uniquely a point P^* on V^*. For this reason the join V^* is often referred to in the literature as the *variety of pairs of corresponding points of V and V'*.

If K is not algebraically closed then P^* need not be uniquely determined by P and P'. The following is an example[25]. Let K be the field of real numbers and let $\Sigma = K(x, y)$, where x and y are indeterminates. We take as V and V' two planes given—in nonhomogeneous coordinates—by the general points (x, y_1) and (x_1, y), respectively, where $y_1 = y(x^2+1)$ and $x_1 = x(y^2+1)$. Here we have: $\mathfrak{o} = K[x, y_1]$, $\mathfrak{o}' = K[x_1, y]$, $\mathfrak{o}^* = (\mathfrak{o}, \mathfrak{o}') = K[x, y]$. Let $\mathfrak{p} = (x^2+1, y_1)$, $\mathfrak{p}' = (x_1, y^2+1)$. These ideals are prime and zero-dimensional in their respective rings and they represent corresponding points of the two planes V and V'. However, the ideal $\mathfrak{o}^*\cdot(\mathfrak{p}, \mathfrak{p}')$ is now the intersection of the following prime ideals: $\mathfrak{p}_1^* = (x^2+1, y-x)$, $\mathfrak{p}_2^* = (x^2+1, y+x)$.

The join V^* of two locally normal varieties V and V' need not be locally normal. One may often find it convenient to pass from V^* to a derived normal variety $\bar{V}^*$ of V^*. We may call $\bar{V}^*$ the *normal join* of V and V'.

---

(25) It should not be too difficult to find necessary and sufficient conditions in order that a given pair of corresponding points P, P' determine uniquely a point P^* of the join V^*. The following is a sufficient condition. *If $\Delta = \mathfrak{o}/\mathfrak{p}$ and $\Delta' = \mathfrak{o}'/\mathfrak{p}'$ are the residue fields of the points P and P' respectively, then a least field Δ^*/K containing Δ/K and Δ'/K should exist such that its relative degree over K is the product of the relative degrees of Δ and of Δ' over K.* (It is not difficult to see that this product is the maximum value for the relative degree of Δ^* over K, and if that maximum is reached, then there exists, to within relative isomorphisms, only one least field which contains Δ and Δ'.) However, the above condition is not sufficient. In fact, no condition can be both necessary and sufficient which does not take into account the quotient rings $Q(P)$ and $Q(P')$ themselves, besides the residue fields Δ and Δ'.

The usefulness of the join V^* is due to the possibility of deriving properties of the birational correspondence between V and V' by first passing from V to V^* and then from V^* to V'. In each of these two steps we are dealing with a birational correspondence between two varieties *which has no fundamental elements on one of the varieties* (*on* V^*). Birational correspondences of this sort are easier to handle, and they in fact play an important role in the general theory and in applications.

6. **Further properties of fundamental varieties.**

THEOREM 12. *Given an irreducible subvariety W of V there exists an algebraic subvariety of V' which we shall denote by $T[W]$ and which has the following properties*:

A. *Each irreducible component of $T[W]$ corresponds to W.*

B. *Each irreducible subvariety W' of V' which corresponds to W lies on $T[W]$.*

The variety $T[W]$ shall be referred to in the sequel as the *transform of W*.

Proof. Suppose that the theorem is true for V and the join V^* of V and V'. Then we show that it is also true for V and V'. For let T^* denote, as before, the birational correspondence between V and V^* and let $T^*[W]=W_1^*+W_2^*+\cdots+W_h^*$, where each W_i^* is irreducible. To each W_i^* there corresponds on V' a unique irreducible variety W_i'. Since, by hypothesis, $T^*(W)=W_i^*$, it follows that each of the varieties $W_1', W_2', \cdots, W_h'$ corresponds to W. On the other hand, let W' be any irreducible subvariety of V' which corresponds to W, and let W^* be an irreducible subvariety of V^* which corresponds to both W and W'. By hypothesis, $W^*\subseteq T^*(W)$, say $W^*\subseteq W_1^*$. Passing to the corresponding subvarieties W' and W_1' of V', we conclude that $W'\subseteq W_1'$, that is, $W'\subseteq W_1'+W_2'+\cdots+W_h'$. Hence $T[W]=W_1'+W_2'+\cdots+W_h'$, where, of course, some of the h varieties W_i' may be embedded, so that the number of irreducible components of $T[W]$ may actually be less than h.

We now prove the theorem for V and for the join V^*. We use nonhomogeneous coordinates. As nonhomogeneous coordinates for V^* we can use the coordinates ξ_{ij} of the preceding section, since the $(n+1)(m+1)$ systems of coordinates $\eta_{ij}/\eta_{\alpha\beta}$ (α and β are fixed for each system) cover the entire projective space in which V^* is embedded. To prove the existence of the transform $T^*[W]$, it will be sufficient therefore to exhibit that part L^* of $T^*[W]$ which is at finite distance with respect to the coordinates ξ_{ij}. Let $\mathfrak{o}$, $\mathfrak{o}'$ and $\mathfrak{o}^*$ have the same meaning as in the preceding section. If W is not at finite distance with respect to the coordinates ξ_i, then *no* W^* which corresponds to W on V^* can be at finite distance with respect to the coordinates ξ_{ij} (since $\mathfrak{o}\subseteq\mathfrak{o}^*$). Hence in this case L^* is empty. If W is at finite distance, it is given by a prime ideal $\mathfrak{p}$ in $\mathfrak{o}$. Let $\mathfrak{p}_1^*, \mathfrak{p}_2^*, \cdots, \mathfrak{p}_h^*$ be those *minimal* prime ideals of $\mathfrak{o}^*\cdot\mathfrak{p}$ which contract to $\mathfrak{p}$, and let W_i^* be the irreducible subvariety of V^*

which is defined by $\mathfrak{p}_i^*$. I assert that $L^* = W_1^* + W_2^* + \cdots + W_k^*$. For in the first place, each W_i^* corresponds to W. In the second place, if $W^* = T^*(W)$ and if W^* is at finite distance with respect to the coordinates ξ_{ij}, W^* is given by a prime $\mathfrak{o}^*$-ideal $\mathfrak{p}^*$, such that $\mathfrak{p}^* \cap \mathfrak{o} = \mathfrak{p}$. Since $\mathfrak{p} \subseteq \mathfrak{p}^*$, we have $\mathfrak{o}^*\mathfrak{p} \subseteq \mathfrak{p}^*$, whence $\mathfrak{p}^*$ must divide some minimal ideal $\mathfrak{p}'^*$ of $\mathfrak{o}^*\mathfrak{p}$. Since $\mathfrak{p} \subseteq \mathfrak{p}'^* \subseteq \mathfrak{p}^*$, it follows that $\mathfrak{p} \subseteq \mathfrak{p}'^* \cap \mathfrak{o} \subseteq \mathfrak{p}$, that is $\mathfrak{p}'^* \cap \mathfrak{o} = \mathfrak{p}$. Therefore $\mathfrak{p}'^*$ is one of the ideals $\mathfrak{p}_1^*, \mathfrak{p}_2^*, \cdots, \mathfrak{p}_k^*$, and since $\mathfrak{p}^* \supseteq \mathfrak{p}'^*$, we conclude that $W^* \subseteq W_1^* + W_2^* + \cdots + W_k^*$. This completes the proof.

COROLLARY 1. *If the birational correspondence between V and V' has no fundamental elements on V', then to a fundamental variety W on V there corresponds on V' at least one variety of higher dimension than W; in other words, $T[W]$ is of higher dimension than W.*

For if there are no fundamental elements on V' and if $W' = T(W)$ then dimension $W' \geqq$ dimension W (Theorem 8, (B) and (C), II.3). If every W' which corresponds to W were of the same dimension as W, then $T[W]$ would be pure ρ-dimensional, where $\rho =$ dimension W. But then the irreducible components of $T[W]$ would be the only subvarieties of V' which correspond to W, and since the number of these components is finite, the corollary follows, by Theorem 10, II.3.

As a consequence, we have the following *characterization of fundamental varieties*:

COROLLARY 2. *If V and V' are arbitrary birationally equivalent varieties and if T and T^* denote, respectively, the birational correspondence between V and V' and the birational correspondence between V and V^*, then a given irreducible subvariety W of V is fundamental for T if and only if $T^*[W]$ is of higher dimension than W.*

For W is fundamental for T if and only if it is fundamental for T^* (Theorem 11, II.5) and since T^* has no fundamental elements on V^*.

In addition to the transform $T[W]$ we shall also have occasion to consider what we call the *total transform of* W and that we shall denote by $T\{W\}$. By that we mean *the locus of points of V' which correspond to points of W*. That $T\{W\}$ is an algebraic variety is seen as follows. As in the case of $T[W]$, so also here it is sufficient to show that $T^*\{W\}$ is algebraic, where T^* is, as usual, the birational correspondence between V and the join V^* of V and V'. Now if we consider that part L_1^* of $T^*\{W\}$ which is at finite distance with respect to the nonhomogeneous coordinates ξ_{ij} we see immediately that *L_1^* is the algebraic subvariety of V^* which is defined by the ideal $\mathfrak{o}^* \cdot \mathfrak{p}$*. For a point P^* of V^*, at finite distance with respect to the ξ_{ij}'s, corresponds to a point P on W, if and only if the corresponding 0-dimensional prime $\mathfrak{o}^*$-ideal $\mathfrak{p}^*$ satisfies the relation: $\mathfrak{p}^* \cap \mathfrak{o} \supseteq \mathfrak{p}$, that is, if and only if $\mathfrak{p}^* \supseteq \mathfrak{o}^* \cdot \mathfrak{p}$.

The irreducible components of $T^*\{W\}$, at finite distance, correspond to

the minimal prime ideals of $\mathfrak{o}^*\cdot\mathfrak{p}$. We have seen that the irreducible components of $T^*[W]$, at finite distance, correspond to those minimal prime ideals of $\mathfrak{o}^*\mathfrak{p}$ which contract to $\mathfrak{p}$. Hence $T^*[W]$ *lies on* $T^*\{W\}$, and from this follows immediately that *also* $T[W]$ *lies on* $T\{W\}$. This result can also be deduced directly from property B, stated in II.1. In the same fashion one sees immediately that $T\{W\}$ also has the following property: *If* $W_1 \subseteq W$, *and if* $W_1' = T(W_1)$, *then* $W_1' \subseteq T\{W\}$.

We point out explicitly that $T[W]$ *may very well be a proper subvariety* of $T\{W\}$. For instance, if T is a plane Cremona transformation and if W is a curve containing a fundamental point P, then $T\{W\} = T[W] + \Gamma'$, where Γ' is the irregular ("fundamental" in the old terminology) curve which corresponds to the point P. Here $T[W]$ is either a curve (if W is regular) or a fundamental point (if W is irregular)[26]. Quite generally, we have the following theorem:

THEOREM 13. *Any irreducible component of the total transform* $T\{W\}$ *which is not a component of the transform* $T[W]$ *must correspond to a proper subvariety* W_1 *of* W. *Moreover, if* V *is locally normal, then* W_1 *must be a fundamental variety.*

Proof. In the proof of Theorem 12 we have seen that if $T^*[W] = W_1^* + W_2^* + \cdots + W_h^*$ and if W_i' is *the* subvariety of V' which corresponds to W_i^*, then $T[W] = W_1' + W_2' + \cdots + W_h'$. In a similar fashion it is seen immediately that if $T^*\{W\} = W_1^* + W_2^* + \cdots + W_h^* + W_{h+1}^* + \cdots$, then $T\{W\} = W_1' + W_2' + \cdots + W_h' + W_{h+1}' + \cdots$. Now let W_0' be an irreducible component of $T\{W\}$ which does not belong to $T[W]$. Then W_0' must correspond to an irreducible component W_0^* of $T^*\{W\}$ which does not belong to $T^*[W]$. Now by definition of $T^*\{W\}$, each point of W_0^* must correspond to some point of W. Since to a point of V^* there corresponds a unique point of V, it follows that the subvariety W_0 of V which corresponds to W_0^* must lie on W. It must be a *proper* subvariety of W, since $W_0^* \not\subseteq T^*[W]$. Now both W_0' and W_0 correspond to W_0^*, whence they correspond to each other. It remains to prove that W_0 is fundamental for T. If W_0 were not funda-

(26) This example shows therefore that the ideal $\mathfrak{o}^*\mathfrak{p}$ *may very well possess minimal prime ideals which contract in* $\mathfrak{o}$ (not to $\mathfrak{p}$ but) *to proper divisors of* $\mathfrak{p}$. In this connection we wish to correct a statement on p. 135 in Krull's Ergebnisse report *Idealtheorie*. The theorem stated on that page consists of three parts, and in the first part it is asserted, among other things, that $\tilde{\mathfrak{p}} = \mathfrak{p}\cdot\tilde{\mathfrak{J}}$. This (and only this) assertion is incorrect. The rings $\mathfrak{J}$ and $\tilde{\mathfrak{J}}$ play there the role of our rings $\mathfrak{o}$ and $\mathfrak{o}^*$. It is quite true that there is only one prime ideal $\tilde{\mathfrak{p}}$ which lies over $\mathfrak{p}$ (that is, such that $\tilde{\mathfrak{p}} \cap \mathfrak{J} = \mathfrak{p}$), and it is also true that $\mathfrak{p}$ is a minimal prime of the ideal $\tilde{\mathfrak{J}}\cdot\mathfrak{p}$. The equality $\tilde{\mathfrak{J}}_{\tilde{\mathfrak{p}}} = \mathfrak{J}_{\mathfrak{p}}$ shows clearly that the case under consideration corresponds to a *regular* W. Nevertheless, $\tilde{\mathfrak{J}}\cdot\mathfrak{p}$ may have minimal prime ideals other than $\tilde{\mathfrak{p}}$, as was pointed out above. Already the quadratic transformation $x' = x$, $y' = y/x$ may serve very well as a source of simple counterexamples. The formal source of that incorrect statement made in Krull's report is the erroneous assertion made earlier on the same page (line 12) to the effect that if the rank equals m then $L_{\mathfrak{p}}$ is a field.

mental, then by a stronger reason W would not be fundamental, since $Q(W)\supset Q(W_0)$ (see Theorem 8 (E)). But then we would have $T[W_0]\subset T[W]$, a contradiction.

If V is not locally normal, one passes to the derived normal varieties of V and of V^* and the rest of the proof is straightforward.

COROLLARY. *If no subvariety of W is fundamental, and if V is locally normal, then the total transform of W coincides with the transform of W.*

We make one more remark about the transforms $T[W]$ and $T\{W\}$. It is clear that in no case can these varieties be empty. However, it may very well happen that with respect to a given system of nonhomogeneous coordinates either $T[W]$ or even $T\{W\}$ is entirely at infinity. Referring to the join V^* and to the rings $\mathfrak{o}$ and $\mathfrak{o}^*$ considered above, we see that $T^*[W]$ is at infinity *if $\mathfrak{o}^*$ does not contain prime ideals which contract to $\mathfrak{p}$* (in the terminology of Krull: $\mathfrak{p}$ *is lost in* $\mathfrak{o}^*$, see [2, p. 134]). If also $T^*\{W\}$ is entirely at infinity, then *the ideal $\mathfrak{o}^*\cdot\mathfrak{p}$ is the unit ideal.*

7. **The main theorem.** If the birational correspondence between V and V' has no fundamental elements on V', and if W is fundamental for T, then, by Theorem 12, Corollary 1, the transform $T[W]$ has at least one component of higher dimension than W. In the general case, that is, when V is an arbitrary variety, that is the best result one may claim, since it is quite possible for $T[W]$ to possess components which have the same dimension as W. However, in the case in which V *is locally normal at* W we have the following important theorem:

MAIN THEOREM. *If W is an irreducible fundamental variety on V of a birational correspondence T between V and V' and if T has no fundamental elements on V', then—under the assumption that V is locally normal at W—each irreducible component of the transform $T[W]$ is of higher dimension than W.*

COROLLARY. *In the more general case in which T has fundamental elements on both V and V' and under the assumption that W is fundamental for T and that V is locally normal at W, the transform $T^*[W]$ of W on the join V^* of V and V' has the property that all its irreducible components are of higher dimension than W.*

We shall first give the main theorem another formulation which is more directly algebraic. Since V is locally normal at W, it is permissible to replace V by a derived normal variety of V (Theorem 7, Corollary 2, II.2). Hence we assume that V is a normal variety. Since there are no fundamental elements on V', the birational correspondence between V' and V^* is regular (Theorem 11, II.5). Hence it is permissible to replace V' by V^*. When these preparations are carried out then the main theorem expresses a feature of the the relationship between the ideals in the two rings $\mathfrak{o}$ and $\mathfrak{o}^*$ considered in II.5, that is, we have to prove the following theorem:

THEOREM 14. *Let $\mathfrak{o}$ and $\mathfrak{o}^*$ be two finite integral domains with the same quotient field Σ, where we assume that the ring $\mathfrak{o}$ is integrally closed and that it is a subring*[27] *of $\mathfrak{o}^*$. Let $\mathfrak{p}$ be a prime ideal in $\mathfrak{o}$ and let $\mathfrak{p}^*$ be a prime ideal $\mathfrak{o}^*$ which lies over $\mathfrak{p}$, that is, such that $\mathfrak{p}^* \cap \mathfrak{o} = \mathfrak{p}$. If $\mathfrak{p}$ and $\mathfrak{p}^*$ have the same dimension, then*

(1) *either the quotient rings $\mathfrak{o}_{\mathfrak{p}}$ and $\mathfrak{o}^*_{\mathfrak{p}^*}$ coincide*[28] *(and in this case $\mathfrak{p}^*$ is obviously the only prime $\mathfrak{o}^*$-ideal which lies over $\mathfrak{p}$) or*

(2) *$\mathfrak{p}^*$ is not minimal with respect to the property of lying over $\mathfrak{p}$, that is, there exists in $\mathfrak{o}^*$ another prime ideal which also lies over $\mathfrak{p}$ and which is a proper multiple of $\mathfrak{p}^*$.*

The proof of this theorem is rather long and will be developed in this and in the following section. That part of the proof which is contained in this section consists of three steps: (a) a reduction to a simpler special case; (b) a reference to the theorem of Krull which we have already mentioned[26]; (c) a lemma.

(a) Let $\mathfrak{o}'^*$ be the integral closure of $\mathfrak{o}^*$ in Σ. It is well known that over every prime ideal $\mathfrak{p}^*$ in $\mathfrak{o}^*$ there lies at least one prime ideal $\mathfrak{p}'^*$ in $\mathfrak{o}'^*$, and that the number of such ideals $\mathfrak{p}'^*$ is finite. Moreover $\mathfrak{p}^*$ and $\mathfrak{p}'^*$ have the same dimension and it is clear that if $\mathfrak{p}^* \cap \mathfrak{o} = \mathfrak{p}$, then $\mathfrak{o}_{\mathfrak{p}} \subseteq \mathfrak{o}^*_{\mathfrak{p}^*} \subseteq \mathfrak{o}'^*_{\mathfrak{p}'^*}$. From this it follows that if our theorem is true for the pair of rings $\mathfrak{o}$ and $\mathfrak{o}'^*$, then it is also true for $\mathfrak{o}$ and $\mathfrak{o}^*$. *Hence we may assume that $\mathfrak{o}^*$ is integrally closed.*

Since $\mathfrak{o}$ and $\mathfrak{o}^*$ are finite integral domains, $\mathfrak{o}^*$ is a finite ring extension of $\mathfrak{o}$. We prefer to think of $\mathfrak{o}^*$ as a ring obtained from $\mathfrak{o}$ by a *finite* number of simple ring extensions, *each ring extension being followed up by the operation of integral closure.* Let therefore:

$$\mathfrak{o}_1 = \mathfrak{o}[\alpha_1], \ \mathfrak{o}_1' = \text{integral closure of } \mathfrak{o}_1;$$
$$\mathfrak{o}_2 = \mathfrak{o}_1'[\alpha_2], \ \mathfrak{o}_2' = \text{integral closure of } \mathfrak{o}_2;$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\mathfrak{o}_m = \mathfrak{o}_{m-1}'[\alpha_m], \ \mathfrak{o}_m' = \mathfrak{o}^* = \text{integral closure of } \mathfrak{o}_m.$$

Let us assume that our theorem is true when $m=1$. Then we show, by induction with respect to m, that the theorem is true for any value of m. Let $\mathfrak{p}^* \cap \mathfrak{o}_{m-1}' = \mathfrak{p}_{m-1}'$. Since $\mathfrak{p}_{m-1}' \cap \mathfrak{o} = \mathfrak{p}$, we have: dimension $\mathfrak{p}^* \geqq$ dimension $\mathfrak{p}_{m-1}' \geqq$ dimension $\mathfrak{p}$. If $\mathfrak{p}$ and $\mathfrak{p}^*$ have the same dimension, then it follows that also $\mathfrak{p}_{m-1}'$ has the same dimension as $\mathfrak{p}$. Now let us also assume that no proper prime multiple of $\mathfrak{p}^*$ lies over $\mathfrak{p}$. Then no proper prime multiple of $\mathfrak{p}^*$ can lie over $\mathfrak{p}_{m-1}'$, and therefore, by the case $m=1$, we have that *$\mathfrak{o}^*_{\mathfrak{p}^*}$ coincides with the quotient ring of $\mathfrak{p}_{m-1}'$ in $\mathfrak{o}_{m-1}'$*. Consequently, there is a (1, 1) correspondence between the prime multiples of $\mathfrak{p}^*$ in $\mathfrak{o}^*$ and the prime multiples

---

[27] It will be seen from the proof that these assumptions can be weakened as follows: *$\mathfrak{o}$ has a finite degree of transcendency over the ground field and is a finite discrete principal order* (see Krull [2, p. 104]); *$\mathfrak{o}^*$ is a finite ring extension of $\mathfrak{o}$.*

[28] This case arises when W is not fundamental, therefore regular.

of $\mathfrak{p}'_{m-1}$ in $\mathfrak{o}'_{m-1}$. (See I.2.) Therefore it is equally true that no prime multiple of $\mathfrak{p}'_{m-1}$ lies over $\mathfrak{p}$. By our induction, we conclude that *also* $\mathfrak{o}_\mathfrak{p}$ *coincides with the quotient ring of* $\mathfrak{p}'_{m-1}$ *in* $\mathfrak{o}'_{m-1}$. Hence $\mathfrak{o}_\mathfrak{p}=\mathfrak{o}^*_\mathfrak{p}$, as was asserted.

We therefore have only to prove our theorem in the following special case: $\mathfrak{o}^*$ *is the integral closure of a ring* $\mathfrak{o}'$, *where* $\mathfrak{o}'$ *is a simple ring extension of* $\mathfrak{o}$: $\mathfrak{o}'=\mathfrak{o}[\alpha]$.

(b) In this special case we shall make use of a theorem stated in Krull [2, p. 135], to which we have already referred in the preceding section(28). We write the principal fractional ideal $\mathfrak{o}\cdot\alpha$ as a quotient of two integral ideals: $\mathfrak{o}\cdot\alpha=\mathfrak{z}/\mathfrak{n}$, where $\mathfrak{z}$ and $\mathfrak{n}$ are symbolic power products of minimal prime ideals in $\mathfrak{o}$, without common factors. Krull distinguishes three cases: (1) $\mathfrak{n}\equiv 0(\mathfrak{p})$, $\mathfrak{z}\not\equiv 0(\mathfrak{p})$; (2) $\mathfrak{n}\not\equiv 0(\mathfrak{p})$; (3) $\mathfrak{n}\equiv 0(\mathfrak{p})$, $\mathfrak{z}\equiv 0(\mathfrak{p})$.

In the first case the element $1/\alpha$ is a non-unit in $\mathfrak{o}_\mathfrak{p}$ and from this it follows immediately that $\mathfrak{p}$ is lost in $\mathfrak{o}'$, that is, no prime ideal in $\mathfrak{o}'$ contracts to $\mathfrak{p}$. But then $\mathfrak{p}$ is also lost in $\mathfrak{o}^*$, since $\mathfrak{o}^*$ is the integral closure of $\mathfrak{o}'$, and in this case there is nothing to prove.

In the second case α is contained in the quotient ring $\mathfrak{o}_\mathfrak{p}$, whence $\mathfrak{o}^*\subset\mathfrak{o}_\mathfrak{p}$, since $\mathfrak{o}_\mathfrak{p}$ is integrally closed. From this we conclude that the two rings $\mathfrak{o}_\mathfrak{p}$, $\mathfrak{o}^*_\mathfrak{p}$ coincide(29). This is the alternative (1) of the theorem.

The really significant case is the third one. In this case Krull's result is to the effect that $\mathfrak{o}'\cdot\mathfrak{p}$ *is a prime ideal* $\mathfrak{p}'$, *that* $\mathfrak{p}'$ *lies over* $\mathfrak{p}$ *and that the dimension of* $\mathfrak{p}'$ *is one greater than the dimension of* $\mathfrak{p}$. We shall make use of this result.

(c) In addition to the above result which concerns the relationship between the ideal theory in $\mathfrak{o}$ and in $\mathfrak{o}'$, we shall have to make use, in a very essential fashion, of the following property of the conductor $\mathfrak{C}$ of $\mathfrak{o}'$ with respect to $\mathfrak{o}^*$:

LEMMA 6. *Each prime* $\mathfrak{o}^*$*-ideal* $\mathfrak{p}^*$ *of the conductor* $\mathfrak{C}$ *has the property that it contracts in* $\mathfrak{o}$ *to a prime ideal of lower dimension.*

Proof. The lemma implies in particular that $\mathfrak{C}$ has no *zero-dimensional* prime ideals. Let us assume that this particular consequence of the lemma has been established and let us show that then the lemma follows by the usual device of ground field extension.

Let $\mathfrak{p}^*$ be a prime $\mathfrak{o}^*$-ideal and let $\mathfrak{p}^*\cap\mathfrak{o}=\mathfrak{p}$. We assume that $\mathfrak{p}$ and $\mathfrak{p}^*$ have the same dimension, say dimension s. It shall now be shown that $\mathfrak{p}^*$ cannot be a prime ideal of $\mathfrak{C}$.

We select in $\mathfrak{o}$ a set of s elements $\zeta_1, \zeta_2, \cdots, \zeta_s$ which are algebraically independent modulo $\mathfrak{p}$. We adjoin these elements to the ground field K getting a new ground field $K_1=K(\zeta_1, \zeta_2, \cdots, \zeta_s)$ and also the new rings: $\mathfrak{o}_1=K_1\cdot\mathfrak{o}$, $\mathfrak{o}'_1=K_1\cdot\mathfrak{o}'$, $\mathfrak{o}^*_1=K_1\cdot\mathfrak{o}^*$. We point out that $\mathfrak{o}^*_1$ is a quotient ring of $\mathfrak{o}^*$, namely $\mathfrak{o}^*_1=\mathfrak{o}^*_S$, where S is the set of all polynomials in $\zeta_1, \zeta_2, \cdots, \zeta_s$, with coeffi-

(29) The proof is exactly the same as the proof which in II.2 led us to the conclusion that the quotient ring of $\mathfrak{m}^*_i$ in the ring $\mathfrak{J}^*$ coincides with $\mathfrak{J}'_i$.

cients in K. Similarly we have: $\mathfrak{o}_1=\mathfrak{o}_S$. We therefore can apply the properties of the correspondence between the ideals in a given ring R and a quotient ring R_S, as described in I.1. We find then that $\mathfrak{p}_1=\mathfrak{o}_1\cdot\mathfrak{p}$ and $\mathfrak{p}_1^*=\mathfrak{o}_1^*\cdot\mathfrak{p}^*$ are prime ideals and $\mathfrak{C}_1=K_1\mathfrak{C}$ is the conductor of $\mathfrak{o}_1'$ with respect to $\mathfrak{o}_1^*$. We have $\mathfrak{o}_1'=\mathfrak{o}_1[\alpha]$ and it is clear that $\mathfrak{o}_1^*$ is the integral closure of $\mathfrak{o}_1'$. Since $\mathfrak{p}_1^*$ is *zero-dimensional* over the ground field K_1, it follows, by our assumption, that $\mathfrak{C}_1:\mathfrak{p}_1^*=\mathfrak{C}_1$. On the other hand we have $\mathfrak{C}_1:\mathfrak{p}_1^*=K_1\cdot(\mathfrak{C}:\mathfrak{p}^*)$. Hence *the two ideals* $\mathfrak{C}$ *and* $\mathfrak{C}:\mathfrak{p}^*$ *have the same extended ideal in* $\mathfrak{o}_1^*$. It follows (see II.1(8)) that these two ideals can only differ by primary components whose associate prime ideals contain polynomials in $\zeta_1, \zeta_2, \cdots, \zeta_s$. Since the ζ's are algebraically independent modulo $\mathfrak{p}^*$, we conclude that $\mathfrak{p}^*$ is not among the prime ideals of $\mathfrak{C}$, as was asserted.

The proof of the lemma is thus reduced to the matter of proving that the conductor $\mathfrak{C}$ does not possess zero-dimensional prime ideals.

Let $\mathfrak{p}^*$ be a zero-dimensional prime ideal in $\mathfrak{o}^*$. We have to prove that $\mathfrak{C}:\mathfrak{p}^*=\mathfrak{C}$. Let ζ^* be an arbitrary element of $\mathfrak{C}:\mathfrak{p}^*$. We denote by $f(x)$ the irreducible polynomial in $K[x]$ such that $f(\alpha)\equiv 0(\mathfrak{p}^*)$. We have then: $\zeta^*f(\alpha)\equiv 0(\mathfrak{C})$, whence $\zeta^*f(\alpha)\cdot\eta^*\in\mathfrak{o}'$, *for any element* η^* *in* $\mathfrak{o}^*$. Hence we may write

$$\zeta^*\eta^*\cdot f(\alpha) = G(\alpha) = \omega_0\alpha^\nu + \omega_1\alpha^{\nu-1} + \cdots + \omega_\nu, \qquad \omega_i\in\mathfrak{o}. \tag{7}$$

We divide through $G(x)$ by $f(x)$:

$$G(x) = A(x)f(x) + R(x),$$

where all polynomials are in $\mathfrak{o}[x]$ and where $R(x)$ is of degree at most $m-1$, if m is the degree of $f(x)$. We now rewrite (7) as follows (notice that $f(\alpha)\neq 0$, since α is not in $\mathfrak{o}$ and since $\mathfrak{o}$ is integrally closed):

$$\zeta^*\eta^* = A(\alpha) + \frac{R(\alpha)}{f(\alpha)}\cdot \tag{8}$$

Let v be an arbitrary valuation of Σ whose valuation ring R_v contains $\mathfrak{o}$. If $v(\alpha)\geqq 0$, then $\mathfrak{o}[\alpha]\in R_v$, and also $\mathfrak{o}^*\subset R_v$. Hence, by (8), we have $R(\alpha)/f(\alpha)\in R_v$. If $v(\alpha)<0$, then $v(R(\alpha))>v(f(\alpha))$, since R is of less degree than f and since $v(f(\alpha))=mv(\alpha)$ (the leading coefficient of $f(\alpha)$ is an element of K). Hence also in this case $R(\alpha)/f(\alpha)$ is contained in R_v. Since this holds true for any valuation v such that $\mathfrak{o}\subset R_v$ and since $\mathfrak{o}$ is integrally closed, we conclude that $R(\alpha)/f(\alpha)\in\mathfrak{o}$. Hence, by (8), $\zeta^*\eta^*\in\mathfrak{o}'$, *for any element* η^* *in* $\mathfrak{o}^*$. Consequently $\zeta^*\in\mathfrak{C}$. Since ζ^* was an arbitrary element of $\mathfrak{C}:\mathfrak{p}^*$, it follows that $\mathfrak{C}:\mathfrak{p}^*=\mathfrak{C}$. This completes the proof of the lemma.

8. Continuation of the proof of the main theorem. To prove the main theorem, or better, the equivalent Theorem 14, we shall proceed as follows. We assume that we have the special case described in the preceding section under (a). Let $\mathfrak{p}^*$ be a prime ideal in $\mathfrak{o}^*$ and let $\mathfrak{p}^*\cap\mathfrak{o}=\mathfrak{p}$. We shall also assume

that we are dealing with the significant case $\mathfrak{n} \equiv 0(\mathfrak{p})$, $\mathfrak{z} \equiv 0(\mathfrak{p})$, in which case we have the result of Krull as stated in the preceding section under (b). We shall prove that if $\mathfrak{p}$ and $\mathfrak{p}^*$ have the same dimension, then $\mathfrak{p}^*$ contains properly another prime ideal $\mathfrak{p}_1^*$ with the property: $\mathfrak{p}_1^* \cap \mathfrak{o} = \mathfrak{p}$. This is the second alternative of Theorem 14.

We divide the proof into two parts, according as $\mathfrak{C} \not\equiv 0(\mathfrak{p}^*)$ or $\mathfrak{C} \equiv 0(\mathfrak{p}^*)$.

First case: $\mathfrak{C} \not\equiv 0(\mathfrak{p}^*)$. Let $\mathfrak{p}^* \cap \mathfrak{o}' = \mathfrak{p}_1'$. Since $\mathfrak{C} \not\equiv 0(\mathfrak{p}_1')$, it follows in an elementary fashion from the very definition of the conductor, that $\mathfrak{p}^*$ is the only prime ideal in $\mathfrak{o}^*$ which contracts to $\mathfrak{p}_1'$ and that the quotient rings $\mathfrak{o}^*_{\mathfrak{p}^*}$, $\mathfrak{o}'_{\mathfrak{p}'_1}$ coincide. We have $\mathfrak{p}_1' \cap \mathfrak{o} = \mathfrak{p}$, whence $\mathfrak{p}_1'$ must be a divisor of the prime ideal $\mathfrak{p}' = \mathfrak{o}' \cdot \mathfrak{p}$. It must be *a proper divisor* of $\mathfrak{p}'$, since $\mathfrak{p}_1'$ is of the same dimension as $\mathfrak{p}$, while by Krull's result $\mathfrak{p}'$ is of dimension one greater than $\mathfrak{p}$. Now since $\mathfrak{C} \not\equiv 0(\mathfrak{p}_1')$, we have *a fortiori* $\mathfrak{C} \not\equiv 0(\mathfrak{p}')$. Hence there is a unique prime ideal $\mathfrak{p}'^*$ in $\mathfrak{o}^*$ which contracts to $\mathfrak{p}'$ and we have $\mathfrak{o}^*_{\mathfrak{p}'^*} = \mathfrak{o}'_{\mathfrak{p}'}$. Since we have also $\mathfrak{o}^*_{\mathfrak{p}^*} = \mathfrak{o}'_{\mathfrak{p}'_1}$ and since $\mathfrak{p}_1'$ is a proper divisor of $\mathfrak{p}'$, it follows that also $\mathfrak{p}^*$ *is a proper divisor of* $\mathfrak{p}'^*$. Since $\mathfrak{p}'^* \cap \mathfrak{o} = \mathfrak{p}' \cap \mathfrak{o} = \mathfrak{p}$, our proof is complete.

Second case: $\mathfrak{C} \equiv 0(\mathfrak{p}^*)$. In this case $\mathfrak{p}^*$ is either a prime ideal of the conductor $\mathfrak{C}$ or properly contains a prime ideal of $\mathfrak{C}$. Since, by hypothesis, $\mathfrak{p}^*$ contracts in $\mathfrak{o}$ to an ideal $\mathfrak{p}$ of the same dimension as $\mathfrak{p}^*$, the first possibility is excluded by our lemma. Hence $\mathfrak{p}^*$ properly contains a prime ideal $\mathfrak{p}_1^*$ of $\mathfrak{C}$. Let $\mathfrak{p}_1^* \cap \mathfrak{o} = \mathfrak{p}_1$ and let

$$\text{dimension } \mathfrak{p} = s, \qquad \text{dimension } \mathfrak{p}_1 = s_1.$$

Since $\mathfrak{p}^* \supset \mathfrak{p}_1^*$, we have $\mathfrak{p} \supseteq \mathfrak{p}_1$. If $\mathfrak{p}_1 = \mathfrak{p}$ then the theorem is proved, since we have now a proper multiple $\mathfrak{p}_1^*$ of $\mathfrak{p}^*$ which also contracts to $\mathfrak{p}$. We assume therefore that $\mathfrak{p} \supset \mathfrak{p}_1$, whence $s_1 > s$. By our lemma, $\mathfrak{p}_1^*$ is of greater dimension than $\mathfrak{p}_1$; by Krull's theorem, the dimension of $\mathfrak{p}_1^*$ is at most one greater than the dimension of $\mathfrak{p}_1$. Hence the dimension of $\mathfrak{p}_1^*$ is s_1+1. Our proof would be complete if we could show that there exists a prime ideal in $\mathfrak{o}^*$, *between* $\mathfrak{p}^*$ and $\mathfrak{p}_1^*$ (and different from $\mathfrak{p}^*$) which contracts to $\mathfrak{p}$. This we proceed to show.

We pass to the residue class rings $\mathfrak{O} = \mathfrak{o}/\mathfrak{p}_1$, $\mathfrak{O}^* = \mathfrak{o}^*/\mathfrak{p}_1^*$. Both rings are finite integral domains and $\mathfrak{O}$ is a subring of $\mathfrak{O}^*$. The first ring is of degree of transcendency s_1, while $\mathfrak{O}^*$ is of degree of transcendency s_1+1. In the homomorphisms $\mathfrak{o} \sim \mathfrak{O}$, $\mathfrak{o}^* \sim \mathfrak{O}^*$, the prime ideals $\mathfrak{p}$ and $\mathfrak{p}^*$ are mapped, respectively, onto prime ideals $\mathfrak{P}$ and $\mathfrak{P}^*$, of the same dimension s, $s < s_1$, and we have: $\mathfrak{P}^* \cap \mathfrak{O} = \mathfrak{P}$. What we have to prove is the existence in $\mathfrak{O}^*$ of a prime ideal which is a proper multiple of $\mathfrak{P}^*$ and which contracts in $\mathfrak{O}$ to the ideal $\mathfrak{P}$. The assertion that such an ideal exists is equivalent to the assertion that $\mathfrak{P}^*$ *is not a minimal prime of the extended ideal* $\mathfrak{O}^* \cdot \mathfrak{P}$. To prove this we pass to the quotient rings $\mathfrak{J} = \mathfrak{O}_{\mathfrak{P}}$, $\mathfrak{J}^* = \mathfrak{O}^*_{\mathfrak{P}^*}$. If $\mathfrak{m}$ and $\mathfrak{m}^*$ denote the ideals of non-units in these two rings, then our problem is to prove that $\mathfrak{m}^*$ is not a minimal prime ideal of the extended ideal $\mathfrak{J}^* \cdot \mathfrak{m}$. The proof of this will complete the proof of the main theorem.

Since $\mathfrak{J}$ is of degree of transcendency s_1, we can find s_1-s elements $\zeta_1, \zeta_2, \cdots, \zeta_{s_1-s}$ in $\mathfrak{J}$ such that the ideal $\mathfrak{A}=\mathfrak{J}\cdot(\zeta_1, \zeta_2, \cdots, \zeta_{s_1-s})$ be exactly s-dimensional. Since the ideal of non-units $\mathfrak{m}$ in $\mathfrak{J}$ is also s-dimensional, it follows that $\mathfrak{A}$ will then be necessarily a primary ideal, with $\mathfrak{m}$ as associated prime. Now consider the ideal $\mathfrak{A}^*=\mathfrak{J}^*\cdot(\zeta_1, \zeta_2, \cdots, \zeta_{s_1-s})$. Since $\mathfrak{J}^*$ is of degree of transcendency s_1+1 and since $\mathfrak{A}^*$ is not the unit ideal (since $\zeta_i\in\mathfrak{A}\equiv 0(\mathfrak{m})$, whence $\mathfrak{A}^*\equiv 0(\mathfrak{m}^*)$), every minimal prime of $\mathfrak{A}^*$ is of dimension at least $s+1$. Let $\mathfrak{P}^*$ be a minimal prime of $\mathfrak{A}^*$. Since $\mathfrak{A}^*\equiv 0(\mathfrak{P}^*)$, we have $\mathfrak{P}^*\cap\mathfrak{J}\supseteq\mathfrak{A}$, whence $\mathfrak{P}^*\cap\mathfrak{J}=\mathfrak{m}$, for $\mathfrak{A}$ is primary and its associated prime is the ideal $\mathfrak{m}$ of non-units. Hence $\mathfrak{P}^*\supseteq\mathfrak{J}^*\mathfrak{m}$, and this shows that $\mathfrak{m}^*$ is not a minimal prime of $\mathfrak{J}^*\cdot\mathfrak{m}$, since $\mathfrak{P}^*$ is a proper multiple of $\mathfrak{m}^*$ (dimension $\mathfrak{m}^*=s$, dimension $\mathfrak{P}^*\geqq s+1$), q.e.d.

In the main theorem we have assumed that the birational correspondence has no fundamental elements on V'. In the general case of an arbitrary pair of birationally equivalent varieties V and V' we may apply the main theorem to V and to the join V^* of V and V'. If we then take into account Theorem 11 of II.5 we deduce the following corollary which expresses the local character of the main theorem:

COROLLARY. *If W is an irreducible subvariety of V at which V is locally normal and if W is fundamental for the birational correspondence T between V and some other variety V', then each irreducible component of $T[W]$ which is not fundamental for T^{-1} is of higher dimension than W.*

9. **The fundamental locus of a birational correspondence.** The forms $\phi_0(\eta), \phi_1(\eta), \cdots, \phi_m(\eta)$ which are proportional to the coordinates $\eta_0', \eta_1', \cdots, \eta_m'$ of the general point of V' (see equations (4), II.5) define a linear system of forms:

$$\phi_\lambda = \lambda_0\phi_0 + \lambda_1\phi_1 + \cdots + \lambda_m\phi_m.$$

We shall allow the parameters λ_i to take arbitrary values (not all zero) in the relative algebraic closure K$'$ of K in Σ, that is, the λ's shall be elements of Σ which are either in K or algebraic over K. We shall also assume that V and V' are locally normal varieties. Under this assumption it is permissible to identify K with K$'$, since any ring of nonhomogeneous coordinates of the general point of a locally normal variety is integrally closed and consequently contains K$'$. We therefore assume that K itself is algebraically closed in Σ.

The principal ideal (ϕ_λ) in the ring K $[\eta_0, \eta_1, \cdots, \eta_m]$ is $(r-1)$-dimensional. Since V is locally normal, the conductor of this ring with respect to its integral closure is a primary irrelevant ideal (or the unit ideal). Hence, to within an irrelevant component which we shall disregard, the ideal (ϕ_λ) is quasi-gleich to a product of symbolic powers of minimal prime (homogeneous) ideals. In particular, let

(9) $$(\phi_i) = \mathfrak{M}\mathfrak{A}_i, \qquad i = 0, 1, \cdots, m,$$

where $\mathfrak{A}_0, \mathfrak{A}_1, \cdots, \mathfrak{A}_m$ have no common factor. Then $\mathfrak{M}$ will be the h.c.d. of all principal ideals (ϕ_λ), that is, we will have: $(\phi_\lambda) = \mathfrak{M}\cdot\mathfrak{A}_{(\lambda)}$. The ideal $\mathfrak{A}_{(\lambda)}$ defines a pure $(r-1)$-dimensional subvariety $\mathfrak{C}_{(\lambda)}$ of V, which may be reducible and in which each irreducible component is counted to a definite multiplicity (equal to the exponent of the corresponding prime factor of $\mathfrak{A}_{(\lambda)}$). As the λ's vary in K, the variety $C_{(\lambda)}$ varies and describes a *linear system* $|C|$ *of* $(r-1)$-*dimensional varieties on* V, free from fixed components since $\mathfrak{A}_0, \mathfrak{A}_1, \cdots, \mathfrak{A}_m$ have no common factor. We have in particular the members $C_0, C_1, \cdots, C_m$ of $|C|$ which correspond to the ideals $\mathfrak{A}_0, \mathfrak{A}_1, \cdots, \mathfrak{A}_m$.

Let F be the algebraic subvariety of V defined by the ideal $\mathfrak{F} = (\mathfrak{A}_0, \mathfrak{A}_1, \cdots, \mathfrak{A}_m)$, or rather, by the radical of this ideal. The variety F is of dimension at most $r-2$ and is common to $C_0, C_1, \cdots, C_m$. We show that *F is the base manifold of the linear system* $|C|$, that is, that *F lies on each* $C_{(\lambda)}$. (This is not obvious, because of the presence of the factor $\mathfrak{M}$.)

Let $\mathfrak{p}$ be a minimal ideal of $\mathfrak{F}$ and let us show that the assumption $\mathfrak{A}_{(\lambda)} \not\equiv 0(\mathfrak{p})$ leads to a contradiction. Let β be an element of $\mathfrak{A}_{(\lambda)}$ not in $\mathfrak{p}$. Since $\phi_i \equiv 0(\mathfrak{M})$, it follows([30]), by (9), that $\beta\phi_i \equiv 0(\phi_{(\lambda)})$, $i = 0, 1, \cdots, m$. Let $\beta\phi_i = \beta_i\phi_{(\lambda)}$, $\beta_i \in K[\eta]$. This relation can be written as follows: $\beta\mathfrak{A}_i = \beta_i\mathfrak{A}_{(\lambda)}$. By hypothesis $\mathfrak{A}_i \equiv 0(\mathfrak{p})$, but $\mathfrak{A}_{(\lambda)} \not\equiv 0(\mathfrak{p})$. Hence $\beta_i \equiv 0(\mathfrak{p})$, and therefore the relations $\beta\phi_i = \beta_i\phi_{(\lambda)}$ yield relations of the form

$$\beta\phi_i = \beta_{i0}\phi_0 + \beta_{i1}\phi_1 + \cdots + \beta_{im}\phi_m, \qquad i = 0, 1, \cdots, m,$$

where $\beta_{ij} \equiv 0(\mathfrak{p})$. From this we conclude that the determinant $\Delta = |\beta_{ij} - \delta_{ij}\beta|$ vanishes ($\delta_{ij} = 0$ if $i \neq j$, $\delta_{ii} = 1$), and this is impossible since $\Delta \equiv \pm\beta^m(\mathfrak{p})$, whence $\Delta \not\equiv 0(\mathfrak{p})$.

THEOREM 15. *The base manifold F of the linear system $|C|$ is also the fundamental locus of the birational transformation T, that is, F has the property that any irreducible subvariety W of V which is fundamental for T lies on F, and conversely.*

Proof. Let $\mathfrak{p}$ be the homogeneous prime ideal in the ring $K[\eta]$ which corresponds to W.

Assume $W \not\subseteq F$. Then at least one of the $m+1$ varieties C_i does not contain F. Let, say, $F \not\subseteq C_0$, whence $\mathfrak{A}_0 \not\equiv 0(\mathfrak{p})$. We introduce the nonhomogeneous coordinates $\xi_i' = \eta_i'/\eta_0'$ of the general point of V' and we denote, as usual,

([30]) Strictly speaking, the congruences $\beta \equiv 0(\mathfrak{A}_{(\lambda)})$, $\phi_i \equiv 0(\mathfrak{M})$ do not necessarily imply that $\beta\phi_i$ is a multiple of $\phi_{(\lambda)}$, since the equation (9) is only true to within an irrelevant component. However, in view of the regularity and the (1, 1) character of the birational correspondence between a locally normal variety and its derived normal variety, it is permissible to give the proof under the assumption that V and V' are not only locally normal but normal. If V is normal, then the equation (9) is exact. The same remark applies to the other theorems proved in this section.

the ring $K[\xi_1', \xi_2', \cdots, \xi_m']$ by $\mathfrak{o}'$. We have $(\xi_i') = \mathfrak{A}_i/\mathfrak{A}_0$, whence $\xi_i' \in Q(W)$. Thus the entire ring $\mathfrak{o}'$ is contained in the quotient ring $Q(W)$, and this implies that *W is not fundamental.*

Conversely, *assume that W is not fundamental.* There corresponds then to W a unique variety W' on V' and we have $Q(W') \subseteq Q(W)$. We may assume that $\eta_0' \neq 0$ on W', and then we will have $\mathfrak{o}' \subset Q(W')$, whence *a fortiori* $\mathfrak{o}' \subset Q(W)$. The fact that the quotients ϕ_i/ϕ_0 all belong to $Q(W)$ leads immediately to the conclusion that $\mathfrak{A}_0 \not\equiv 0(\mathfrak{p})$. *Hence W does not lie on F,* q.e.d.

While Theorem 15 gives full information about the location of the fundamental elements of a birational correspondence, the following theorem tells us where the *irregular* varieties are located:

THEOREM 16. *If an irreducible subvariety W' of V' is irregular for T^{-1} then W' lies on the total transform $T\{F\}$ of the fundamental locus F of T. Conversely, if W' lies on $T\{F\}$, then it is either irregular or fundamental for T^{-1}.*

The proof is immediate. For if W' is irregular, and if $W = T^{-1}(W')$, then W is fundamental (Theorem 8 (C), II.3), whence $W \subseteq F$. Consequently, $W' = T(W) \subseteq T\{F\}$. Conversely, if $W' \subseteq T\{F\}$, then W' must correspond to some irreducible subvariety W of F. Since W is fundamental, W' cannot be regular, q.e.d.

The linear system $|C|$ which is defined by the linear family of forms $\phi_{(\lambda)}$ has always played an important part in the study of the birational correspondence T with which this system is associated. Theorem 15 is one illustration of the geometric connection between $|C|$ and T. Another property of the linear system $|C|$ which follows in a straightforward fashion from the definition is the following: *the birational correspondence T transforms the linear system $|C|$ into the system of hyperplane sections of V'.* This statement should be intended in the following sense: *a general* member $C_{(\lambda)}$ of $|C|$ is regular and $T(C_{(\lambda)})$ is the section $\Gamma'_{(\lambda)}$ of V' with the hyperplane $\lambda_0 y_0' + \lambda_1 y_1' + \cdots + \lambda_m y_m' = 0$. However, for special values of the λ's, it may very well happen that $C_{(\lambda)}$, or $\Gamma'_{(\lambda)}$, or both, contain irreducible components which are irregular and which therefore correspond to fundamental varieties.

10. **Isolated fundamental varieties.** We assume as before that V is locally normal and we keep the notation V^* for the join of V and V', and T^* for the birational correspondence between V and V^*.

DEFINITION 8. *Let $W_1^*, W_2^*, \cdots, W_h^*$ be the irreducible components of the total transform $T^*\{F\}$, where F is the fundamental locus of T^* (and hence also of T) on V. The irreducible subvarieties $F_i = T^{*-1}(W_i^*)$ of V are called the isolated fundamental varieties of T (and also of T^*).*

It is clear that the isolated fundamental varieties F_i lie on the fundamental locus F. It is also not difficult to see that *the irreducible components of F are among the isolated fundamental varieties.* For let W be an irreducible compo-

nent of F and let W^* be an irreducible component of the transform $T^*[W]$. We have $W^* \subseteq T^*\{F\}$, since $T^{*-1}(W^*) = W \subseteq F$. Consequently W^* lies on one of the varieties W_i^*, say $W^* \subseteq W_1^*$. We have then $T^{*-1}(W^*) \subseteq T^{*-1}(W_1^*) = F_1$, that is, $W \subseteq F_1$. Since W is a component of F, we conclude that $F_1 = W$.

It is important to point out that in addition to the irreducible components of the fundamental locus F there may exist other "embedded" isolated fundamental varieties, which are proper subvarieties of the irreducible components of F. Thus in the three-dimensional case we may have a fundamental curve Γ on V to which there corresponds a surface on V^*, and on that fundamental curve Γ there may exist some special point P to which there also corresponds a surface on V^*. This point P must be regarded as an isolated fundamental point, although it is embedded in the fundamental curve Γ. The term "isolated" refers not to the position of the point P with respect to the fundamental locus F but to its role in the birational correspondence T.

By the main theorem each irreducible component W_i^* of $T^*\{F\}$ is of higher dimension than the corresponding isolated fundamental variety F_i. Under certain conditions it is possible to assert that *each* W_i^* *is of dimension* $r-1$. We proceed to find such conditions.

Let η_{ij} denote as usual the homogeneous coordinates of the general point of the join V^*, where $\eta_{ij} = \eta_i \phi_j$ (see II.5 and II.9). Let us consider quite generally an arbitrary homogeneous ideal $(g_0, g_1, \cdots, g_h)$ in the ring $K[\eta_0, \eta_1, \cdots, \eta_n]$, where each g_i is a form, say of degree ν_i. We put

$$g_{ij} = g_i(\eta_{0j}, \eta_{1j}, \cdots, \eta_{nj}), \qquad i = 0, 1, \cdots, h; j = 0, 1, \cdots, m. \tag{10}$$

The $(m+1)(h+1)$ forms g_{ij} in the η_{ij}'s generate a homogeneous ideal in the ring $K[\eta_{00}, \eta_{10}, \cdots, \eta_{nm}]$.

LEMMA 7. *If N is the subvariety of V defined by the ideal $(g_0, g_1, \cdots, g_h)$ and if N^* is the subvariety of V^* defined by the ideal $(g_{00}, \cdots, g_{hm})$, then N^* is the total transform of N, that is, $N^* = T^*\{N\}$.*

Proof. Let W and W^* be two corresponding irreducible subvarieties of V and V^*, respectively. We have to show that $W^* \subseteq N^*$, if and only if $W \subseteq N$.

Assume that $W \subseteq N$. Without loss of generality we may assume that $\eta_0 \neq 0$ on W. Then if v denotes a valuation of centers W and W^*, we will have:

$$v(\eta_i/\eta_0) \geqq 0, \qquad v(g_i/\eta_0^{\nu_i}) > 0. \tag{11}$$

Also without loss of generality we may assume that $v(\phi_i/\phi_0) \geqq 0$, for $i = 1, 2, \cdots, m$. We will have then $v(\eta_{ij}/\eta_{00}) = v(\eta_i/\eta_0) + v(\phi_j/\phi_0) \geqq 0$. By (10), we can write:

$$g_{ij} = g_i \phi_j^{\nu_i}, \tag{12}$$

and hence $g_{ij}/\eta_{00}^{\nu_i} = g_i/\eta_0^{\nu_i} \cdot (\phi_j/\phi_0)^{\nu_i}$. Consequently $v(g_{ij}/\eta_{00}^{\nu_i}) > 0$, by (11), and this shows that $W^* \subseteq N^*$.

Conversely, assume that $W^* \subseteq N^*$. Then if $\eta_{00} \neq 0$ on W^*, we find that $\eta_0 \neq 0$ on W (since $\eta_i/\eta_0 = \eta_{i0}/\eta_{00}$). On the other hand we have: $g_{i0}/\eta_{00}^{\nu_i} = g_i/\eta_0^{\nu_i}$, whence $v(g_i/\eta_0^{\nu_i}) > 0$, that is, $W \subseteq N$, q.e.d.

We now apply the lemma to the case in which N is given by the ideal $(\phi_0, \phi_1, \cdots, \phi_m)$. Then N^* is given by the ideal $(\phi_{00}, \phi_{10}, \cdots, \phi_{mm})$, where

$$\phi_{ij} = \phi_i(\eta_{0j}, \eta_{1j}, \cdots, \eta_{nj}).$$

The relations (12) now yield: $\phi_{ij} = \phi_i \phi_j^{\nu}$, where ν is the common degree of the form ϕ_i. From these relations we deduce the following: $\phi_{ij}^{\nu+1} = \phi_{ii}\phi_{jj}^{\nu}$, and consequently the two ideals $(\phi_{00}, \phi_{10}, \cdots, \phi_{mm})$ and $(\phi_{00}, \phi_{11}, \cdots, \phi_{mm})$ have the same radical. Therefore N^* is also defined by the ideal $(\phi_{00}, \phi_{11}, \cdots, \phi_{mm})$. Now we have $\phi_{ii}/\phi_{jj} = (\phi_i/\phi_j)^{\nu+1} = \eta_{ki}^{\nu+1}/\eta_{kj}^{\nu+1}$, for any $i, j = 0, 1, \cdots, m$; $k = 0, 1, \cdots, n$. Therefore each irreducible component of N^* at which $\eta_{kj} \neq 0$, for some k and j, is also a component of the principal ideal (ϕ_{jj}) and is therefore $(r-1)$-dimensional. *Consequently N^* is pure $(r-1)$-dimensional.*

In view of the formulas (9) of II.9, the variety of the ideal $(\phi_0, \phi_1, \cdots, \phi_m)$ consists of the $(r-1)$-dimensional variety of the ideal $\mathfrak{M}$ and of the fundamental locus F. We therefore can assert that *$T^*\{F\}$ is pure $(r-1)$-dimensional if $\mathfrak{M}$ is the unit ideal.* The hypothesis $\mathfrak{M} = (1)$ implies that *each member $C_{(\lambda)}$ of the linear system $|C|$ associated with the birational correspondence is complete intersection of V with a hypersurface of the ambient projective space,* namely with the hypersurface $\lambda_0\phi_0(y_0, y_1, \cdots, y_n) + \lambda_1\phi_1(y_0, y_1, \cdots, y_n) + \cdots + \lambda_m\phi_m(y_0, y_1, \cdots, y_n) = 0$.

Conversely, let us assume that each $C_{(\lambda)}$ is *complete intersection*. Then in particular C_0 is complete intersection, whence the ideal $\mathfrak{A}_0$ is a principal ideal, say $\mathfrak{A}_0 = (\psi_0)$. We have: $\phi_i\psi_0/\phi_0 = \mathfrak{M}\mathfrak{A}_i \cdot \mathfrak{A}_0/\mathfrak{M}\mathfrak{A}_0 = \mathfrak{A}_i$, that is, $\phi_i\psi_0/\phi_0$ is an integral ideal. Consequently the quotients $\phi_i\psi_0/\phi_0$ are forms in the η's, say $\phi_i\psi_0/\phi_0 = \psi_i$. The forms $\psi_0, \psi_1, \cdots, \psi_m$ are proportional to the forms $\phi_0, \phi_1, \cdots, \phi_m$, and the linear system $|C|$ is also defined by the linear family of forms $\lambda_0\psi_0 + \lambda_1\psi_1 + \cdots + \lambda_m\psi_m$. If we use the ψ's instead of the ϕ's, we will have $\mathfrak{M} = (1)$, and we reach again the conclusion that $T^*\{F\}$ is pure $(r-1)$-dimensional.

We can go a step further. Let us point out that if we define a projective model V'_ρ by the condition that the homogeneous coordinates of its general point be given by a linear base of the forms of degree ρ in $\eta'_0, \eta'_1, \cdots, \eta'_m$ then V' and V'_ρ are in regular birational correspondence (Lemma 5, II.2). The transition from V' to V'_ρ is equivalent to passing from the linear system $|C|$ to the least linear system which contains as members all sets of ρ C's. Hence, by the preceding result, we conclude that *if a sufficiently high multiple of a C is complete intersection, then $T^*\{F\}$ is pure $(r-1)$-dimensional.*

In order to conclude with a similar result of a local character, let F_1 be any isolated fundamental variety of T, and let us assume that C is *locally, at F_1,* complete intersection. By that we mean that some hypersurface cuts V

along C and along a residual variety *which does not contain* F_1. If we replace the ϕ's by a suitable set of proportional forms, we may arrange matters so that the variety of the ideal $\mathfrak{M}$ does not contain F_1. Since $N=M+F$, it is clear that $T^*\{N\}=T^*[M]+T^*\{F\}$, where we write $T^*[M]$ instead of $T^*\{M\}$, since $T^*\{M\}-T^*[M]$ lies on $T^*\{F\}$ (Theorem 13, II.6). Since $F_1 \not\subseteq M$, no component of $T^*[F_1]$ can lie on $T^*[M]$. It follows that the irreducible components of $T^*\{F\}$ which correspond to the isolated fundamental variety F_1 are also components of $T^*\{N\}$, and hence are $(r-1)$-dimensional. The same conclusion is reached if we assume that some sufficiently high multiple of C is complete intersection locally at F_1.

The above results refer to V and to the join of V and V'. In particular, if the birational correspondence T has no fundamental elements on V' then V' may play the role of V^*, since V' and V^* are then in regular birational correspondence. We reassume our results in the following theorem:

THEOREM 17. *If a birational correspondence T between two locally normal r-dimensional varieties V and V' has no fundamental elements on V' and if F denotes the fundamental locus of T on V, then an irreducible component of $T\{F\}$ is of dimension $r-1$, provided the corresponding isolated fundamental variety F_1 has the property that the members of the linear system $|C|$ associated with T, or their sufficiently high multiples, are complete intersections locally, at F_1.*

COROLLARY. *To an isolated simple fundamental variety there always corresponds an $(r-1)$-dimensional variety on V'* (see van der Waerden [6, p. 154]).

For locally, at a simple subvariety of V, every $(r-1)$-dimensional subvariety of V is complete intersection([31]).

11. **Monoidal transformations.** Given a homogeneous ideal $\mathfrak{A}$ in the ring $K[\eta_0, \eta_1, \cdots, \eta_n]$ of homogeneous coordinates of the general point of V, it is possible to associate with $\mathfrak{A}$ an infinite set S of birational transforms of V such that: (1) the birational correspondence between V and any variety V' of the set has no fundamental elements on V' and such that (2) any two varieties of the set are in regular birational correspondence. The varieties V' of the set S shall be defined as follows. Let us take a base of $\mathfrak{A}$ consisting of forms of least possible degrees, and let a be the highest degree of the forms in that base. We define V' by its general point $(\phi_0, \phi_1, \cdots, \phi_m)$, where the ϕ's form a linear base for the forms of a given degree ν in $\mathfrak{A}$ and where we impose on ν the condition: $\nu \geqq a+1$. For $\nu=a+1, a+2, \cdots$, we get an infinite set of models V', and this is our set S. We shall denote these models by $V'_{a+1}, V'_{a+2}, \cdots$.

([31]) For the case of algebraically closed ground fields of characteristic zero see our paper [8, p. 664]. There the proof is given explicitly for surfaces only, but actually exactly the same proof applies to higher varieties. For ground fields which are not algebraically closed or which are of characteristic p, the statement can be derived from the following result obtained by Irvin Cohen in his dissertation: if the characteristic of a complete p-series ring coincides with the characteristic of its residue field, the ring is a power series ring over a field.

First of all it is clear that each V_ν' *is birationally equivalent to* V. For $\mathfrak{A}$ contains at least one form ψ of degree $\nu-1$ so that the products $\eta_0\psi, \eta_1\psi, \cdots, \eta_n\psi$ can be identified with $n+1$ of the ϕ's. This shows that the quotients ϕ_i/ϕ_0 generate the field Σ.

Since $\mathfrak{A}$ has a basis consisting of forms of degree at most a, it follows that if $(\phi_0, \phi_1, \cdots, \phi_m)$ is a basis for the forms in $\mathfrak{A}$ of degree ν, then the products $\eta_i\phi_j$ constitute a basis for the forms in $\mathfrak{A}$ which are of degree $\nu+1$. This holds true also for $\nu=a$. From this it follows that $V_{\nu+1}$ is the join of V and V_ν, provided $\nu \geqq a+1$. Therefore the birational correspondence between V and V_ν has no fundamental points on V_ν, provided $\nu \geqq a+2$. But then $V_{\nu+1}$, *the join of* V *and* V_ν, *is a regular birational transform of* V_ν, always provided that $\nu \geqq a+2$. As for the case $\nu=a+1$, we can still regard V_{a+1} as the join of V and V_a, although in this case V_a need not be birationally equivalent to V. At any rate, the proof that the birational correspondence between V and V_{a+1} has no fundamental elements on V_{a+1} is exactly the same as that given for the join in II.5.

Thus we may say that a given homogeneous ideal $\mathfrak{A}$ in the ring $K[\eta_0, \eta_1, \cdots, \eta_n]$ determines, to within a regular birational transformation, a birational transform V' of V such that the birational correspondence between V and V' has no fundamental elements on V'.

Let N be the subvariety of V defined by the ideal $\mathfrak{A}$. It is quite clear that if $\nu \geqq a$, then the ideal generated by our base $(\phi_0, \phi_1, \cdots, \phi_m)$ differs from $\mathfrak{A}$ only by an irrelevant component. Hence, by Theorem 15 of II.9, we conclude that *if* N *is of the dimension at most* $r-2$, *then* N *is the fundamental locus* F *of the birational correspondence between* V *and* V'. In particular, if N *is empty*, that is, if $\mathfrak{A}$ *is an irrelevant ideal, then* V' *is a regular transform of* V.

If, however, N is of dimension $r-1$, then the fundamental locus F will consist of the irreducible components of N which are of dimension less than $r-1$ and possibly of some proper subvarieties of the $(r-1)$-dimensional components of N. Thus, even in the case in which N is pure $(r-1)$-dimensional, it may very well happen that F is not empty. According to Theorem 15, this will happen if the residual intersections of the hypersurfaces $\phi_i=0$ $(i=0, 1, \cdots, m)$ with V, outside N, have a base manifold on N.

Let now $\mathfrak{A}_1$ be another homogeneous ideal in the ring $K[\eta_0, \eta_1, \cdots, \eta_n]$. If W is an irreducible subvariety of V given by a prime ideal $\mathfrak{p}$, we shall say that $\mathfrak{A}$ and $\mathfrak{A}_1$ *coincide locally* at W, if the two ideals differ only by primary components whose associated prime ideals are not multiples of $\mathfrak{p}$. In other words, $\mathfrak{A}$ and $\mathfrak{A}_1$ coincide locally at W if they give rise to one and the same ideal in the quotient ring of W (see I.1).

LEMMA 8. *If* $\mathfrak{A}$ *and* $\mathfrak{A}_1$ *coincide locally at an irreducible subvariety* W *of* V *and if* V' *and* V_1' *are the birational transforms of* V *which are determined (to within a regular birational transformation), respectively, by* $\mathfrak{A}$ *and by* $\mathfrak{A}_1$, *then*

any irreducible subvariety of V' which corresponds to W is regular for the birational correspondence between V' and V_1'.

Proof. Let $\mathfrak{A}^*$ be the ideal which is obtained from either $\mathfrak{A}$ or $\mathfrak{A}_1$ by the omission of all primary components whose associated prime ideals are not multiples of $\mathfrak{p}$. Let V^* be the birational transform of V determined by the ideal $\mathfrak{A}^*$. It is sufficient to prove the lemma for $\mathfrak{A}$ and $\mathfrak{A}^*$, and for $\mathfrak{A}_1$ and $\mathfrak{A}^*$. We shall prove it, for instance, for $\mathfrak{A}$ and $\mathfrak{A}^*$.

Let W' be an irreducible subvariety of V' which corresponds to W. We have to prove that W' is regular for the birational correspondence between V' and V^*. Let $\phi_0, \phi_1, \cdots, \phi_m$ be a linear base of the forms of degree ν which belong to $\mathfrak{A}$. Since $\mathfrak{A}\subseteq\mathfrak{A}^*$, we may complete this base to a linear base $\phi_0, \phi_1, \cdots, \phi_m, \phi_{m+1}, \cdots, \phi_{m+\mu}$ for the forms of degree ν which belong to the ideal $\mathfrak{A}^*$. We take ν sufficiently high, so that $(\phi_0, \phi_1, \cdots, \phi_m)$ and $(\phi_0, \phi_1, \cdots, \phi_m, \phi_{m+1}, \cdots, \phi_{m+\mu})$ are the general points, respectively, of V' and V^*.

Let v be any valuation whose center on V' is W' and whose center on V is W. We may assume that $v(\phi_i/\phi_0)\geqq 0$, $i=1, 2, \cdots, m$ whence W' is at finite distance with respect to the nonhomogeneous coordinates $\xi_i' =\phi_i/\phi_0$, $i=1, 2, \cdots, m$, of the general point of V'. Let W^* be the center of v on V^*. By our definition of the ideal $\mathfrak{A}^*$, there exists a form $g(\eta_0, \eta_1, \cdots, \eta_n)$ such that $g\cdot\mathfrak{A}^*\equiv 0(\mathfrak{A})$ and such that $g\neq 0$ on W. We have then, in particular: $g\phi_{m+j}=A_0\phi_0+A_1\phi_1+\cdots+A_m\phi_m$, $j=1, 2, \cdots, \mu$, where $A_0, A_1, \cdots, A_m$ are forms in $\eta_0, \eta_1, \cdots, \eta_n$, of the same degree as g. We write:

$$\phi_{m+j}/\phi_0 = \frac{A_0}{g}+\frac{A_1}{g}\cdot\frac{\phi_1}{\phi_0}+\cdots+\frac{A_m}{g}\cdot\frac{\phi_m}{\phi_0}\cdot \tag{13}$$

Since $g\neq 0$ on W we have $v(A_i/g)\geqq 0$, $i=0, 1, \cdots, m$. Since also $v(\phi_i/\phi_0)\geqq 0$, it follows from the above relation (13) that $v(\phi_{m+j}/\phi_0)\geqq 0$. Hence W^* is at finite distance with respect to the nonhomogeneous coordinates $\xi_1', \xi_2', \cdots, \xi_{m+\mu}'$ of the general point of V^*, where $\xi_i' =\phi_i/\phi_0$. Since the ring $K[\xi_1', \xi_2', \cdots, \xi_m']$ is a subring of the ring $K[\xi_1', \xi_2', \cdots, \xi_{m+\mu}']$, it follows that $Q(W')\subseteq Q(W^*)$.

On the other hand, since $g\neq 0$ on W, the quotients A_i/g belong to the quotient ring $Q(W)$. Since $Q(W)\subseteq Q(W')$ and since also the quotients ϕ_i/ϕ_0, $i=1, 2, \cdots, m$, are in $Q(W')$, we conclude, by (13), that the entire ring $K[\xi_1', \xi_2', \cdots, \xi_{m+\mu}']$ is contained in $Q(W')$. From this it follows immediately that $Q(W^*)\subseteq Q(W')$, whence $Q(W^*)=Q(W')$, q.e.d.

Corollary. *If $\mathfrak{A}$ and $\mathfrak{A}_1$ differ only by an irrelevant primary component, then V' and V_1' are in regular birational correspondence.*

From the above general consideration we pass to the special case which interests us, namely to the case in which the given homogeneous ideal $\mathfrak{A}$ is a

prime ideal $\mathfrak{p}$, of dimension s, $0 \leqq s \leqq r-2$. Let W be the irreducible subvariety of V defined by $\mathfrak{p}$. The birational transformation T determined by the ideal $\mathfrak{p}$ (that is, by a linear base of forms of sufficiently high degree in $\mathfrak{p}$) is called a *monoidal transformation of center* W. In the special case when W is a point P the transformation is called *quadratic* (of center P). The birational transform V' of V, under a monoidal transformation of given center, is determined to within a regular birational correspondence. The center W of a monoidal transformation is the fundamental locus of the transformation. Moreover, from Theorem 17, II.10, it follows that in the present case $T\{W\}$ is pure $(r-1)$-dimensional. However, it should be pointed out that $T\{W\}$ may very well be reducible and—this is significant—*some components of* $T\{W\}$ *may correspond to proper subvarieties of* W. In other words, *the center* W *of a monoidal transformation is not necessarily the only isolated fundamental variety of the transformation*(32). We shall see presently that this complication arises only if W carries some singular points of V or if W itself has singularities.

Of special importance in applications are monoidal transformations with *simple center*, that is, with center at a simple subvariety W of V. The special case of a quadratic transformation with simple center has been considered in our paper [11]. The results established there carry over to monoidal transformations with simple center, in view of the following considerations. Let the ground field K be extended by the adjunction of s elements of $Q(W)$ which are algebraically independent on W. With respect to this new ground field K_1, the variety W becomes a (simple) point and the monoidal transformation T becomes a quadratic transformation. Therefore certain properties of the monoidal transformation T, over K, can be deduced from corresponding properties of the quadratic transformation over K_1. However, only such properties of T can be deduced in this fashion as concern W *as a whole*. What happens to special points or special subvarieties of W requires new considerations. For instance, we have proved in the quoted paper [11] that if T is a quadratic transformation with simple center P, then the transform $T[P]$ (which, since P is a point, coincides with the total transform $T\{P\}$) is an irreducible, simple and $(r-1)$-dimensional subvariety of V' and, moreover, that every

(32) Here is an example. Let V be the quadric hypersurface $u^2 = yz$ in the 4-dimensional space of the variables x, y, z, u. This hypersurface has the double line $y=z=u=0$. Let W be the line $x=y=u=0$. As nonhomogeneous coordinates of the monoidal transform V' of V we can take the elements $x, y, z, u, x/u, y/u$. Let $\mathfrak{o}'$ be the ring of these coordinates and let $\mathfrak{o} = K[x, y, z, u]$. We have $\mathfrak{o}' = \mathfrak{o}[x/u, y/u] = K[x_1, y_1, z]$, where $x_1 = x/u$, $y_1 = y/u$. Here $\mathfrak{p} = (x, y, u)$ is the prime ideal of W and we have $\mathfrak{o}' \cdot \mathfrak{p} = \mathfrak{o}' \cdot u = \mathfrak{o}' \cdot y_1 z$, that is, that part of $T\{W\}$ which is at finite distance consists of two planes: $y_1 = 0$ and $z = 0$ (note that the *affine* model V' is in regular birational correspondence with the affine space of the variables x_1, y_1, z). The first plane corresponds to W (since $\mathfrak{o}' \cdot y_1 \cap \mathfrak{o} = \mathfrak{p}$). But the plane $z=0$ corresponds to the point $x=y=z=u=0$. This point is imbedded in W, but according to our terminology must be regarded as an isolated fundamental point.

point of $T[P]$ is likewise simple for V'. Now when we pass from the ground field K to the ground field K_1, we lose all those components of $T\{W\}$ which cannot be regarded as varieties over K_1, that is, all those components of $T\{W\}$ which correspond to proper subvarieties of W (since on any proper subvariety of W the s elements which have been adjointed to K are algebraically dependent). Consequently, the correct extrapolation of the above result concerning quadratic transformations to monoidal transformations is the following:

THEOREM 18. *If the center W of a monoidal transformation T is a simple subvariety of V, then the transform*(33) *$T[W]$ of W is an irreducible, simple, $(r-1)$-dimensional subvariety of V', and every irreducible subvariety W' of $T[W]$ is also simple for V', provided $W'=T(W)$.*

The total transform $T\{W\}$ may possess components which are not components of $T[W]$ (even if W is simple(32)), and concerning those components we can assert nothing. Likewise $T[W]$ may contain points which are singular for V. Thus, if V is three-dimensional and if W is a curve, then $T[W]$ is a surface which may carry, in addition to a finite number of singular points of V, also a finite number of singular curves of V, but each such curve must correspond to a point of W.

The following theorem will show, among other things, that these complications can arise only from points or subvarieties of W which are singular for V or for W.

THEOREM 19. *Let W_1 be an irreducible subvariety of W, of dimension s_1. If W_1 is simple both for V and W, then $T[W_1]$ lies on $T[W]$, is irreducible, is of dimension $r-1-s+s_1$ and is simple both for V' and for $T[W]$. Moreover, every irreducible subvariety of $T[W_1]$ which corresponds to W_1 is likewise simple for V', $T[W]$ and also for $T[W_1]$.*

Proof. By the usual device of ground field extension we can achieve a reduction to the case $s_1=0$. Therefore we assume that W_1 is a point P of W, simple both for V and W. It is then possible to select uniformizing parameters $t_1, t_2, \cdots, t_r$ at P in such a fashion that W be locally, at P, complete intersection of the $r-s$ hypersurfaces(34) $t_1=0$, $t_2=0, \cdots, t_{r-s}=0$. Then

(33) Not the *total* transform $T\{W\}$!

(34) **Proof.** Quite generally, the uniformizing parameters $t_1, t_2, \cdots, t_{r-\rho}$ of a simple ρ-dimensional subvariety L of V have the following property: *if $g(t_1, t_2, \cdots, t_{r-\rho})=0$ is a true homogeneous relation between these parameters, with coefficients in the quotient ring $Q(L)$, then all these coefficients must be zero on L, that is, they are non-units of $Q(L)$.* (See [9, p. 202, (15) and p. 207, (23)].) In view of this property and also because $Q(L)$ is a chain theorem ring in which the non-units form an ideal, the quotient ring of a simple subvariety is a p-series ring (*p-Reihenring*) in the sense of Krull [3]. We shall therefore apply properties of p-series rings due to Krull.

Let $\mathfrak{J}$ denote the quotient ring of P and let $\mathfrak{m}$ be the ideal of non-units in $\mathfrak{J}$. If $\tau_1, \tau_2, \cdots, \tau_r$ are uniformizing parameters of P, then we have $\mathfrak{m}=\mathfrak{J}\cdot(\tau_1, \tau_2, \cdots, \tau_r)$. If α is any element of $\mathfrak{J}$ and if $\alpha\equiv 0(\mathfrak{m}^h)$, $\alpha\not\equiv 0(\mathfrak{m}^{h+1})$, then α can be written as a form $g_h(\tau_1, \tau_2, \cdots, \tau_r)$, of degree h,

$t_1, t_2, \cdots, t_{r-s}$ will be uniformizing parameters for W (that is, the ideal generated by $t_1, t_2, \cdots, t_{r-s}$ in $Q(W)$ will be the prime ideal of non-units), and the ideal generated by the same elements in $Q(P)$ will be the prime ideal of W in $Q(P)$.

The uniformizing parameters $t_1, t_2, \cdots, t_{r-s}$ of W are proportional to certain forms $\psi_1, \psi_2, \cdots, \psi_{r-s}$ in the homogeneous coordinates $\eta_0, \eta_1, \cdots, \eta_n$ of the general point of V. Since these uniformizing parameters belong to $Q(P)$, it follows that the factor of proportionality can be so selected that the ideal generated by the forms $(\psi_1, \psi_2, \cdots, \psi_{r-s})$ coincide locally at P with the prime ideal of W. Hence by Lemma 8 we can replace the transformation T by the birational transformation defined by the ideal(35) $(\psi_1, \psi_2, \cdots, \psi_{r-s})$. Therefore we may assume that T, instead of being our original monoidal transformation of center W, is the birational transformation which carries V into the variety V' whose general point is $(\eta_{01}, \eta_{11}, \cdots, \eta_{n,r-s})$, where(36) $\eta_{ij}=\eta_i\psi_j$.

Without loss of generality we may assume that the point P (and hence also W) is at finite distance with respect to the nonhomogeneous coordinates $\xi_i=\eta_i/\eta_0$, $i=1, 2, \cdots, n$. Let Γ' denote an irreducible component of $T[P]$. For some value of h, $h=1, 2, \cdots, r-s$, it will be true that Γ' is at finite distance with respect to the nonhomogeneous coordinates

$$\xi_{ij}^{(h)} = \eta_i\psi_j/\eta_0\psi_h$$

---

with coefficients in $\mathfrak{J}$. If the coefficients of this form are replaced by their residues mod $\mathfrak{m}$, one obtains a form in $\tau_1, \tau_2, \cdots, \tau_r$ with coefficients in the residue field $\mathfrak{J}/\mathfrak{m}$. The property of uniformizing parameters stated above implies that this form is uniquely determined by the element α. This form is called by Krull the *leading form* of α [3, p. 207].

It is a straightforward matter to show that r elements $t_1, t_2, \cdots, t_r$ are uniformizing parameters of P, that is, $\mathfrak{m}=\mathfrak{J}\cdot(t_1, t_2, \cdots, t_r)$, if and only if the leading forms of $t_1, t_2, \cdots, t_r$ are linear and linearly independent.

Let $\mathfrak{p}$ denote the prime ideal of W in $\mathfrak{J}$ and let $\mathfrak{J}^*=\mathfrak{J}/\mathfrak{p}$, $\mathfrak{m}^*=\mathfrak{m}/\mathfrak{p}$. Then $\mathfrak{J}^*$ is the quotient ring of the point P, *regarded as a point of* W, and $\mathfrak{m}^*$ is the ideal of non-units of $\mathfrak{J}^*$. Since, by hypothesis, P is a simple point of W, there exist s elements in $\mathfrak{J}^*$, say $t^*_{r-s+1}, t^*_{r-s+2}, \cdots, t^*_r$ such that $\mathfrak{m}^*=\mathfrak{J}^*\cdot(t^*_{r-s+1}, t^*_{r-s+2}, \cdots, t^*_r)$. Let $t_{r-s+1}, t_{r-s+2}, \cdots, t_r$ be elements of $\mathfrak{J}$ whose $\mathfrak{p}$-residues are $t^*_{r-s+1}, t^*_{r-s+2}, \cdots, t^*_r$, respectively. We will have then: $\mathfrak{m}=\mathfrak{J}\cdot(\mathfrak{p}, t_{r-s+1}, t_{r-s+2}, \cdots, t_r)$. From this relation we draw the following consequences. In the first place it follows that the ideal $\mathfrak{p}$ must contain $r-s$ elements whose leading forms are linear and linearly independent. Let $t_1, t_2, \cdots, t_{r-s}$ be such $r-s$ elements of $\mathfrak{p}$. If $\mathfrak{J}'$ is the perfect closure of $\mathfrak{J}$ (see Krull [3, p. 217]), then it is a straightforward matter to show that the ideal $\mathfrak{J}'\cdot(t_1, t_2, \cdots, t_{r-s})$ is prime. Therefore also the ideal $\mathfrak{J}\cdot(t_1, t_2, \cdots, t_{r-s})$ is prime, since it is the contraction of the ideal $\mathfrak{J}'\cdot(t_1, t_2, \cdots, t_{r-s})$ (Krull [3, Theorem 15]). Since the leading ideal of $\mathfrak{J}\cdot(t_1, t_2, \cdots, t_{r-s})$ is of dimension s, it follows (Krull [3, Theorem 8]) that also $\mathfrak{J}\cdot(t_1, t_2, \cdots, t_{r-s})$ is of dimension s. *Consequently this ideal coincides with* $\mathfrak{p}$. We have therefore: $\mathfrak{m}=\mathfrak{J}\cdot(t_1, t_2, \cdots, t_{r-s}, t_{r-s+1}, \cdots, t_r)$, $\mathfrak{J}\cdot\mathfrak{p}=\mathfrak{J}\cdot(t_1, t_2, \cdots, t_{r-s})$, q.e.d.

(35) Note that if two ideals coincide locally at some W they also coincide locally at any W_1 such that $W\subseteq W_1$.

(36) Since our new transformation behaves locally at P as the given monoidal transformation, we could refer to our new transformation as being *locally monoidal at P*.

of the general point of V'. Let $\mathfrak{o}_h'$ denote the ring of these nonhomogeneous coordinates, and let $\mathfrak{o}$ denote, as usual, the ring $K[\xi_1, \xi_2, \cdots, \xi_n]$. Since $\psi_j/\psi_h = t_j/t_h$, we find:

$$\mathfrak{o}_h' = \mathfrak{o}[t_1/t_h, t_2/t_h, \cdots, t_{r-s}/t_h]. \tag{14}$$

Without loss of generality we may assume that Γ' is at finite distance with respect to the nonhomogeneous coordinates $\xi_{ij}^{(1)}$. For simplicity we shall drop the subscript 1 in the symbol $\mathfrak{o}_1'$, that is, we shall use the symbol $\mathfrak{o}'$ to denote the ring $\mathfrak{o}_1'$.

Let $\mathfrak{p}_0$ denote the prime zero-dimensional $\mathfrak{o}$-ideal of the point P. We shall denote by $\mathfrak{p}$ the prime $\mathfrak{o}$-ideal of W. For clarity of exposition we divide our proof into several steps.

(1) We shall show first that *the ideal* $\mathfrak{o}'\cdot\mathfrak{p}_0$ *is prime*. Let $\mathfrak{J} = \mathfrak{o}_{\mathfrak{p}_0}$ denote the quotient ring of P and let $\mathfrak{J}' = \mathfrak{J}\cdot\mathfrak{o}'$. The ring $\mathfrak{J}'$ is a quotient ring of $\mathfrak{o}'$, namely $\mathfrak{J}' = \mathfrak{o}_S'$ where $S = \mathfrak{o} - \mathfrak{p}_0$. Therefore, in view of the relationship between the ideals in a ring and the ideals in its quotient ring, *the ideal* $\mathfrak{o}'\cdot\mathfrak{p}_0$ *is prime if and only if* $\mathfrak{J}'\cdot\mathfrak{p}_0$ *is prime*([37]). We prefer to deal with the ring $\mathfrak{J}'$ and to show that $\mathfrak{J}'\cdot\mathfrak{p}_0$ *is prime*.

Since $\mathfrak{J}\cdot\mathfrak{p}_0 = \mathfrak{J}\cdot(t_1, t_2, \cdots, t_r)$ and $t_i = t_1\cdot t_i/t_1$, $i = 2, 3, \cdots, r-s$, that is, $t_2, t_3, \cdots, t_{r-s}$ are multiplies of t_1 in $\mathfrak{o}'$, it follows that

$$\mathfrak{J}'\cdot\mathfrak{p}_0 = \mathfrak{J}'\cdot(t_1, t_{r-s+1}, \cdots, t_r). \tag{15}$$

Any element α in $\mathfrak{J}'$ can be written in the form:

$$\alpha = \phi_\rho(t_1, t_2, \cdots, t_{r-s})/t_1^\rho,$$

where ϕ_ρ is a form of degree ρ in $t_1, t_2, \cdots, t_{r-s}$, with coefficients in $\mathfrak{J}$. Let β be another element in $\mathfrak{J}'$,

$$\beta = \psi_\sigma(t_1, t_2, \cdots, t_{r-s})/t_1^\sigma,$$

and let us assume that $\alpha\beta \equiv 0(\mathfrak{J}'\cdot\mathfrak{p}_0)$. We will have a relation of the form:

$$\begin{aligned}
&\phi_\rho(t_1, t_2, \cdots, t_{r-s})\cdot\psi_\sigma(t_1, t_2, \cdots, t_{r-s})\cdot t_1^\mu \\
&\quad = [t_1 f_\mu(t_1, t_2, \cdots, t_{r-s}) + t_{r-s+1}f_\mu^{(1)}(t_1, t_2, \cdots, t_{r-s}) + \cdots \\
&\qquad\qquad + t_r f_\mu^{(s)}(t_1, t_2, \cdots, t_{r-s})]t_1^{\rho+\sigma},
\end{aligned}$$

where $f_\mu, f_\mu^{(1)}, \cdots, f_\mu^{(s)}$ are forms of degree μ, with coefficients in $\mathfrak{J}$. The right-hand side of this relation is a form of degree $\rho+\sigma+\mu+1$ in $t_1, t_2, \cdots, t_r$, with

([37]) We have a (1, 1) isomorphic correspondence between the ideals in $\mathfrak{J}'$ and those ideals in $\mathfrak{o}'$ all prime ideals of which contract in $\mathfrak{o}$ to $\mathfrak{p}_0$ or to multiples of $\mathfrak{p}_0$. Now since $\mathfrak{p}_0$ is a maximal ideal, every prime ideal of $\mathfrak{o}'\cdot\mathfrak{p}_0$ contracts to $\mathfrak{p}_0$. Hence $\mathfrak{o}'\cdot\mathfrak{p}_0$ and $\mathfrak{J}'\cdot\mathfrak{p}_0$ are corresponding ideals in the above correspondence.

coefficients in $\mathfrak{J}$. Hence by a well known property of uniformizing parameters([34]) either all the coefficients of ϕ_ρ or all the coefficients of ψ_σ must be elements of $\mathfrak{J}\cdot\mathfrak{p}_0$. Suppose that all the coefficients of ϕ_ρ are in $\mathfrak{J}\cdot\mathfrak{p}_0$. Since $\mathfrak{J}\cdot\mathfrak{p}_0=\mathfrak{J}\cdot(t_1, t_2, \cdots, t_r)$, we will have for ϕ_ρ an expression of the form: $\phi_\rho=t_1\phi_\rho^{(1)}(t_1,t_2,\cdots,t_{r-s})+t_2\phi_\rho^{(2)}(t_1,t_2,\cdots,t_{r-s})+\cdots+t_r\phi_\rho^{(r)}(t_1,t_2,\cdots,t_{r-s})$, where the ϕ_ρ^j are again forms of degree ρ, with coefficients in $\mathfrak{J}$. Since $t_2, \cdots, t_{r-s}$ are multiples of t_1 in $\mathfrak{J}'$, we conclude immediately that ϕ_ρ/t_1^ρ is contained in the ideal $\mathfrak{J}'(t_1, t_{r-s+1}, \cdots, t_r)$, that is, in view of (15), $\alpha\equiv 0(\mathfrak{J}'\cdot\mathfrak{p}_0)$. This shows that $\mathfrak{J}'\cdot\mathfrak{p}_0$ is a prime ideal, as was asserted([38]).

(2) Let $\mathfrak{o}'\mathfrak{p}_0=\mathfrak{p}'$. *We assert* that *the $\mathfrak{p}'$-residues of $t_2/t_1, \cdots, t_{r-s}/t_1$ are algebraically independent* (over K). For a relation of algebraic dependence between these residues would imply a relation of the form:

$$\phi_\rho(t_1, t_2, \cdots, t_{r-s})/t_1^\rho = \sum_{i=1}^{r} t_i g_\sigma^{(i)}(t_1, t_2, \cdots, t_{r-s})/t_1^\sigma,$$

where the $g_\sigma^{(i)}$ and ϕ_ρ are forms in $t_1, t_2, \cdots, t_{r-s}$, with coefficients in $\mathfrak{o}$, and *where the coefficients of ϕ_ρ are not all in* $\mathfrak{p}_0$. Such a relation, cleared of the denominators, is in contradiction with the property of uniformizing parameters stated above([34]).

From the fact that $\mathfrak{o}'\cdot\mathfrak{p}_0$ is prime, follows that Γ' *is the only irreducible component of $T[P]$ which is at finite distance with respect to the coordinates* $\xi_{ij}^{(1)}$.

The fact that the $\mathfrak{p}'$-residue of $t_2/t_1, \cdots, t_{r-s}/t_1$ are algebraically independent, in conjunction with the fact that $\mathfrak{p}'\cap\mathfrak{o}$ is the *zero-dimensional* ideal $\mathfrak{p}_0$, *implies that Γ' is of dimension $r-s-1$*. Moreover, the algebraic independence of the quotients $t_2/t_1, t_3/t_1, \cdots, t_{r-s}/t_1$ implies in particular that they do not belong to $\mathfrak{p}'$. Hence these quotients are units in the quotient ring $Q(\Gamma')$. But then also $t_i/t_h\in Q(\Gamma')$, for $i, h=1, 2, \cdots, r-s$, whence the rings $\mathfrak{o}_h'$ (see (14)) belong to $Q(\Gamma')$. This shows that Γ' is *at finite distance also with respect to the nonhomogeneous coordinates* $\xi_{ij}^{(h)}$, for $h=1, 2, \cdots, r-s$. Consequently, Γ' is the only irreducible component of $T[P]$, that is, $T[P]$ is irreducible: $T[P]=\Gamma'$.

(3) Let C' denote the irreducible $(r-1)$-dimensional variety $T[W]$. We are interested in the quotient rings of Γ' and of C'. On the basis of the preceding considerations we find immediately that every element of $Q(\Gamma')$ is of the form: $f_\rho(t_1, t_2, \cdots, t_{r-s})/g_\rho(t_1, t_2, \cdots, t_{r-s})$, where f_ρ and g_ρ are forms of like degree ρ, with coefficients in $\mathfrak{o}$, and where the coefficients of g_ρ are not all in $\mathfrak{p}_0$. Similarly, making use of the remark in footnote 38, or also directly from the properties of what we have called "p-adic divisor" in [11], we conclude that the elements of $Q(C')$ are all of the form $f_\rho(t_1, t_2, \cdots, t_{r-s})/g_\rho(t_1, t_2, \cdots, t_{r-s})$,

([38]) Exactly the same proof could be applied toward proving the following: if $\mathfrak{p}$ is the prime deal of W in $\mathfrak{o}$ and if $\mathfrak{J}=Q(W)$, $\mathfrak{J}'=\mathfrak{J}\cdot\mathfrak{o}'$, then the ideal $\mathfrak{J}'\cdot\mathfrak{p}(=\mathfrak{J}'\cdot t_1)$ is prime. From this we could conclude that $T[W]$ is irreducible (as asserted in Theorem 18) in exactly the same fashion as we concluded in the text that $T[P]$ is irreducible.

with the same conditions on f_ρ and g_ρ as above, except that now the coefficients of g_ρ *must not all be* in $\mathfrak{p}$, where $\mathfrak{p}$ is the prime ideal of W. Since $P\subseteq W$, it follows that $Q(\Gamma')\subseteq Q(C')$, that is, $T[P]$ *lies on* $T[W]$ (that is, Γ' lies on C').

Moreover, from the preceding considerations (see relation (15)), it follows that the prime ideal of non-units in $Q(\Gamma')$ has the basis $t_1, t_{r-s+1}, \cdots, t_r$, consisting of $s+1$ elements. Since Γ' is of dimension $r-s-1$, it follows that Γ' *is a simple subvariety of* V', and that $t_1, t_{r-s+1}, \cdots, t_r$ are uniformizing parameters of Γ'. In a similar fashion we find that t_1 is a uniformizing parameter of the $(r-1)$-dimensional variety C', and since t_1 is among the uniformizing parameters of Γ', we conclude that Γ' *is a simple subvariety of* C'.

(4) To complete the proof of our theorem we have only to show that every point P' of Γ' is simple for V', C' and Γ'. To show that P' is simple for V' we have to exhibit r uniformizing parameters at P'. Let Δ be the residue class field of the point P, that is, let $\Delta=\mathfrak{o}/\mathfrak{p}$. Similarly let Δ' be the residue class field of P'. Here Δ and Δ' are finite algebraic extensions of K, and $\Delta'\supseteq\Delta$ since $Q(P')\supseteq Q(P)$. Without loss of generality we may assume that P' is at finite distance with respect to the ring $\mathfrak{o}'$ of nonhomogeneous coordinates ($\mathfrak{o}'=\mathfrak{o}_1'$, see (14)), and is therefore given in $\mathfrak{o}'$ by a prime zero-dimensional ideal $\mathfrak{p}_0'$.

Let $c_1, c_2, \cdots, c_n$ be the P-residues of $\xi_1, \xi_2, \cdots, \xi_n$ respectively ($c_i\in\Delta$) and let $d_2, d_3, \cdots, d_{r-s}$ be the P'-residues of $t_2/t_1, t_3/t_1, \cdots, t_{r-s}/t_1 (d_i\in\Delta')$. The element d_i will be the root of an irreducible polynomial $f_i(z)$ with coefficients in Δ. We replace the coefficients of $f_i(z)$ by arbitrary but fixed elements of $\mathfrak{o}$ of which they are residues. Let $F_i(z)$ be the polynomial with coefficients in $\mathfrak{o}$ thus obtained, $i=2, 3, \cdots, r-s$. *If we assume that the polynomials* $f_1(z)$ *are all separable* then we can conclude as in [11, p. 590] that the r elements

$$(16)\qquad t_1,\ F_2(t_2/t_1),\ \cdots,\ F_{r-s}(t_{r-s}/t_1),\ t_{r-s+1},\ \cdots,\ t_r$$

are uniformizing parameters at P'. Since these elements include the uniformizing parameters $t_1, t_{r-s+1}, \cdots, t_r$ of Γ' and the uniformizing parameter t_1 of C', the proof is complete.

However, if some or all of the polynomials $f_i(z)$ are inseparable, then the elements (16) are no longer uniformizing parameters at P' (compare with footnote 41). We shall therefore give here a new proof which applies both to the separable and non-separable case. We consider the residue class ring $\mathfrak{o}^*=\mathfrak{o}'/\mathfrak{p}'$, where $\mathfrak{p}'$ is the prime $\mathfrak{o}'$-ideal of Γ'. Let $z_2, z_3, \cdots, z_{r-s}$ be the $\mathfrak{p}'$-residues of $t_2/t_1, t_3/t_1, \cdots, t_{r-s}/t_1$. Since $\mathfrak{p}'\cap\mathfrak{o}=\mathfrak{p}_0$ and since we have shown earlier in this section that $z_2, z_3, \cdots, z_{r-s}$ are algebraically independent over K, it follows from (14), for $h=1$, that $\mathfrak{o}^*$ is a *polynomial ring*[39] *over* Δ:

[39] This shows incidentally that the field of rational functions on Γ' is a pure transcendental extension of Δ, whence Γ' *is a rational variety* over Δ.

$$\mathfrak{o}^* = \Delta[z_2, z_3, \cdots, z_{r-s}].$$

The rest of the proof will be based on the following lemma:

LEMMA 9. *In a polynomial ring $P_n = \Delta[x_1, x_2, \cdots, x_n]$ over an arbitrary field Δ every prime zero-dimensional ideal possesses a base consisting of n elements*(40).

Proof of the lemma. Since the lemma is trivially true for $n=1$, we proceed by induction with respect to n. Let $\mathfrak{p}$ be a prime zero-dimensional ideal in P_n and let $f(z)$ be the irreducible polynomial in $\Delta[z]$ such that $f(x_n) \equiv 0(\mathfrak{p})$. The residue class ring $P_{n-1}^* = P_n/f(x_n)$ is obviously a polynomial ring $\Delta^*[x_1, x_2, \cdots, x_{n-1}]$ over the field $\Delta^* = \Delta(\alpha)$, where α is a root of $f(z)$. The ideal $\mathfrak{p}^* = \mathfrak{p}/f(x_n)$ is prime and zero-dimensional in P_{n-1}^*. By our induction, there exist $n-1$ elements $\omega_1^*, \omega_2^*, \cdots, \omega_{n-1}^*$ in P_{n-1}^* such that $\mathfrak{p}^* = P_{n-1}^*(\omega_1^*, \omega_2^*, \cdots, \omega_{n-1}^*)$. Let $\omega_1, \omega_2, \cdots, \omega_{n-1}$ be elements of P_n whose residues modulo $f(x_n)$ are, respectively, $\omega_1^*, \omega_2^*, \cdots, \omega_{n-1}^*$. Then it is clear that $\mathfrak{p} = (\omega_1, \omega_2, \cdots, \omega_{n-1}, f(x_n))$, q.e.d.(41).

We now apply our lemma. In the homomorphism $\mathfrak{o}' \sim \mathfrak{o}^*$ the prime $\mathfrak{o}'$-ideal $\mathfrak{p}_0'$ of the point P' is mapped upon a prime zero-dimensional $\mathfrak{o}^*$-ideal $\mathfrak{p}_0^*$. By the lemma we have

$$\mathfrak{p}_0^* = (\zeta_2^*, \zeta_3^*, \cdots, \zeta_{r-s}^*).$$

Let $\zeta_2', \zeta_3', \cdots, \zeta_{r-s}'$ be elements of $\mathfrak{o}'$ whose $\mathfrak{p}'$-residues are respectively $\zeta_2^*, \zeta_3^*, \cdots, \zeta_{r-s}^*$. Then we have

$$\mathfrak{p}_0' = (\mathfrak{p}', \zeta_2', \zeta_3', \cdots, \zeta_{r-s}'). \tag{17}$$

Let $\mathfrak{J}_1'$ be the quotient ring of P'. Since $Q(P') \supset Q(P) = \mathfrak{J}$ and since $Q(P') \supset \mathfrak{o}'$, the ring $\mathfrak{J}_1'$ contains all the rings previously considered, that is, the rings $\mathfrak{o}$, $\mathfrak{o}'$, $\mathfrak{J}$ and $\mathfrak{J}'$ ($\mathfrak{J}' = \mathfrak{J} \cdot \mathfrak{o}'$). We have $\mathfrak{p}' = \mathfrak{o}'\mathfrak{p}_0$, whence $\mathfrak{J}_1' \cdot \mathfrak{p}' = \mathfrak{J}_1' \cdot \mathfrak{p}_0 = \mathfrak{J}_1' \cdot \mathfrak{J} \cdot \mathfrak{p}_0 = \mathfrak{J}_1' \cdot (t_1, t_2, \cdots, t_r) = \mathfrak{J}_1' \cdot (t_1, t_{r-s+1}, \cdots, t_r)$ (since $t_i/t_1 \in \mathfrak{o}' \in \mathfrak{J}_1'$, $i = 2, 3, \cdots, r-s$). Substituting into (17) we find:

$$\mathfrak{J}_1' \cdot \mathfrak{p}_0' = \mathfrak{J}_1' \cdot (t_1, \zeta_2, \zeta_3, \cdots, \zeta_{r-s}, t_{r-s+1}, t_{r-s+2}, \cdots, t_r).$$

---

(40) This lemma implies that every point P of an affine (or of a projective) space over Δ is simple. The lemma gives, however, a stronger result, since it shows that uniformizing parameters at P can be so selected that they generate the prime ideal of P not only in $Q(P)$ but also in the polynomial ring. In other words: *every point P of an affine n-space is complete intersection of n hypersurfaces.* This result can be extended without any difficulties to projective spaces by a similar inductive argument.

(41) In the case of ground fields of characteristic zero we have used instead of the above lemma the following property of the polynomial ring P_n: if $f_i(z)$ is the irreducible polynomial in $\Delta[z]$ such that $f_i(x_i) \equiv 0(\mathfrak{p})$, then the ideal $(f_1(x_1), f_2(x_2), \cdots, f_n(x_n))$ is the intersection of prime zero-dimensional ideals, one of which is of course the ideal $\mathfrak{p}$ itself. In this case the n polynomials $f_i(x_i)$ are uniformizing parameters for the point defined by $\mathfrak{p}$. This reasoning applies also in the case in which the polynomials $f_i(x_i)$ are separable.

This exhibits r uniformizing parameters at P', and since these parameters include the uniformizing parameters of Γ' and that of C', the proof is now complete.

COROLLARY. *If all points of W are simple both for W and for V, then $T[W]=T\{W\}$ (T—a monoidal transformation of center W) and $T[W]$ is irreducible, $(r-1)$-dimensional and all its points are simple both for $T[W]$ and V'. Moreover, if W is of dimension s, then $T[W]$ is covered by an s-dimensional algebraic system $\{\Gamma'\}$ of $(r-s-1)$-dimensional varieties Γ' in $(1, 1)$ correspondence with the points of W. Each Γ' is irreducible, rational and free from singularities, and through each point of $T[W]$ there passes a unique Γ'.*

REFERENCES

1. H. Grell, *Beziehungen zwischen den Idealen verschiedener Ringe*, Math. Ann. vol. 97 (1927).
2. W. Krull, *Idealtheorie*, Ergebnisse der Mathematik und ihrer Grenzgebiete, IV 3.
3. ———, *Dimensionstheorie in Stellenringen*, J. Reine Angew. Math., vol. 179 (1938).
4. S. MacLane and O. F. G. Schilling, *Zero-dimensional branches of rank one on algebraic varieties*, Ann. of Math. (2) vol. 40 (1939).
5. F. K. Schmidt, *Über die Erhaltung der Kettensätze der Idealtheorie bei beliebigen endlichen Körpererweiterungen*, Math. Zeit. vol. 41 (1936).
6. B. L. van der Waerden, *Algebraische Korrespondenzen und rationale Abbildungen*, Math. Ann. vol. 110 (1934).
7. O. Zariski, *Some results in the arithmetic theory of algebraic varieties*, Amer. J. Math. vol. 61 (1939).
8. ———, *The reduction of the singularities of an algebraic surface*, Ann. of Math. vol. 40 (1939).
9. ———, *Algebraic varieties over ground fields of characteristic zero*, Amer. J. Math. vol. 62 (1940).
10. ———, *Pencils on an algebraic variety and a new proof of a theorem of Bertini*, Trans. Amer. Math. Soc. vol. 50 (1941).
11. ———, *A simplified proof for the resolution of singularities of an algebraic surface*, Ann. of Math. (2) vol. 43 (1942).
12. ———, *Normal varieties and birational correspondences*, Bull. Amer. Math. Soc. vol. 48 (1942).

THE JOHNS HOPKINS UNIVERSITY,
BALTIMORE, MD.

7